

Chapter 3: Linear Functions

3.1 Let $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $\alpha_1, \alpha_2 \in \Re$. Homogeneity implies that

$$f(\alpha_1 \mathbf{x}_1) = \alpha_1 f(x_1)$$
$$f(\alpha_2 \mathbf{x}_2) = \alpha_2 f(x_2)$$

and additivity implies

$$f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2)$$

Conversely, assume

$$f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $\alpha_1, \alpha_2 \in \Re$. Letting $\alpha_1 = \alpha_2 = 1$ implies

$$f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2)$$

while setting $\mathbf{x}_2 = \mathbf{0}$ implies

$$f(\alpha_1 \mathbf{x}_1) = \alpha_1 f(\mathbf{x}_1)$$

3.2 Assume $f_1, f_2 \in L(X, Y)$. Define the mapping $f_1 + f_2 \colon X \to Y$ by

$$(f_1 + f_2)(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$$

We have to confirm that $f_1 + f_2$ is linear, that is

$$\begin{aligned}
(f_1 + f_2)(\mathbf{x}_1 + \mathbf{x}_2) &= f_1(\mathbf{x}_1 + \mathbf{x}_2) + f_2(\mathbf{x}_1 + \mathbf{x}_2) \\
&= f_1(\mathbf{x}_1) + f_1(\mathbf{x}_2) + f_2(\mathbf{x}_1) + f_2(\mathbf{x}_2) \\
&= f_1(\mathbf{x}_1) + f_2(\mathbf{x}_1) + f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) \\
&= (f_1 + f_2)(\mathbf{x}_1) + (f_1 + f_2)(\mathbf{x}_2)
\end{aligned}$$

and

$$\begin{aligned}
(f_1 + f_2)(\alpha \mathbf{x}) &= f_1(\alpha \mathbf{x}) + f_2(\alpha \mathbf{x}) \\
&= \alpha(f_1(\mathbf{x}) + f_2(\mathbf{x})) \\
&= \alpha(f_1 + f_2)(\mathbf{x})
\end{aligned}$$

Similarly let $f \in L(X, Y)$ and define $\alpha f \colon X \to Y$ by

$$(\alpha f)(\mathbf{x}) = \alpha f(\mathbf{x})$$

αf is also linear, since

$$\begin{aligned}
(\alpha f)(\beta \mathbf{x}) &= \alpha f(\beta \mathbf{x}) \\
&= \alpha \beta f(\mathbf{x}) \\
&= \beta \alpha f(\mathbf{x}) \\
&= \beta(\alpha f)(\mathbf{x})
\end{aligned}$$

3.3 Let $\mathbf{x}, \mathbf{x}^1, \mathbf{x}^2 \in \Re^2$. Then

$$\begin{aligned} f(\mathbf{x}^1 + \mathbf{x}^2) &= f(x_1^1 + x_1^2, x_2^1 + x_2^2) \\ &= \big((x_1^1 + x_1^2)\cos\theta - (x_2^1 + x_2^2)\sin\theta, (x_1^1 + x_1^2)\sin\theta - (x_2^1 + x_2^2)\cos\theta\big) \\ &= \big((x_1^1\cos\theta - x_2^1\sin\theta) + (x_1^2\cos\theta - x_2^2\sin\theta), (x_1^1\sin\theta + x_2^1\cos\theta) + (x_1^2\sin\theta - x_2^2\cos\theta)\big) \\ &= \big((x_1^1\cos\theta - x_2^1\sin\theta, x_1^1\sin\theta + x_2^1\cos\theta) + (x_1^2\cos\theta - x_2^2\sin\theta, x_1^2\sin\theta - x_2^2\cos\theta\big) \\ &= f(x_1^1, x_2^1) + f(x_1^2, x_2^2) \\ &= f(\mathbf{x}^1) + f(\mathbf{x}^2) \end{aligned}$$

and

$$\begin{aligned} f(\alpha\mathbf{x}) &= f(\alpha x_1, \alpha x_2) \\ &= (\alpha x_1\cos\theta - \alpha x_2\sin\theta, \alpha x_1\sin\theta + \alpha x_2\cos\theta) \\ &= \alpha\,(x_1\cos\theta - x_1\sin\theta, x_1\sin\theta + x_2\cos\theta) \\ &= \alpha f(x_1, x_2) \\ &= \alpha f(\mathbf{x}) \end{aligned}$$

3.4 Let $\mathbf{x}, \mathbf{x}^1, \mathbf{x}^2 \in \Re^3$.

$$\begin{aligned} f(\mathbf{x}^1 + \mathbf{x}^2) &= f(x_1^1 + x^2, x_2^1 + x_2^2, x_3^1 + x_3^2) \\ &= (x_1^1 + x_1^2, x_2^1 + x_2^2, 0) \\ &= (x_1^1, x_2^1, 0) + (x_1^2, x_2^2, 0) \\ &= f(x_1^1, x_2^1, x_3^1) + f(x_1^2, x_2^2, x_3^2) \\ &= f(\mathbf{x}^1) + f(\mathbf{x}^2) \end{aligned}$$

and

$$\begin{aligned} f(\alpha\mathbf{x}) &= f(\alpha x_1, \alpha x_2, \alpha x_3) \\ &= (\alpha x_1, \alpha x_2, 0) \\ &= \alpha(x_1, x_2, 0) \\ &= \alpha f(x_1, x_2, x_3) \\ &= \alpha f(\mathbf{x}) \end{aligned}$$

This mapping is the projection of 3-dimensional space onto the (2-dimensional) plane.

3.5 Applying the definition

$$\begin{aligned} f(x_1, x_2) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_2, x_1) \end{aligned}$$

This function interchanges the two coordinates of any point in the plane $\Re^2$. Its action corresponds to reflection about the line $x_1 = x_2$ (45 degree diagonal).

3.6 Assume (N, w) and (N, w') are two games in $\mathcal{G}^N$. For any coalition $S \subseteq N$

$$\begin{aligned} (w + w')(S) - (w + w')(S \setminus \{i\}) &= w(S) + w(S') - w(S \setminus \{i\}) - w'(S \setminus \{i\}) \\ &= (w(S) - w(S \setminus \{i\})) + (w'(S) - w'(S \setminus \{i\})) \\ &= \varphi_i(w) + \varphi_i(w') \end{aligned}$$

3.7 The characteristic function of cost allocation game is

$$\begin{array}{lll} w(AP) = 0 & w(AP,TN) = 210 & \\ w(TN) = 0 & w(AP,KM) = 770 & w(N) = 1530 \\ w(KM) = 0 & w(KM,TN) = 1170 & \end{array}$$

The following table details the computation of the Shapley value for player AP.

S	γ_S	$w(S)$	$w(S \setminus \{i\})$	$\gamma_S(w(S) - w(S \setminus \{i\}))$
AP	$1/3$	0	0	0
AP, TN	$1/6$	210	0	35
AP, KM	$1/6$	770	0	$128\,1/3$
AP, TN, KM	$1/3$	1530	1170	120
$\varphi_f(w)$				$283\,1/3$

Thus $\varphi_{AP}w = 283\,1/3$. Similarly, we can calculate that $\varphi_{TN}w = 483\,1/3$ and $\varphi_{KM}w = 763\,1/3$.

3.8

$$\begin{aligned} \sum_{i \in N} \varphi_i w &= \sum_{i \in N} \left(\sum_{S \ni i} \gamma_S (w(S) - w(S \setminus \{i\})) \right) \\ &= \sum_{S \ni i} \left(\sum_{i \in N} \gamma_S (w(S) - w(S \setminus \{i\})) \right) \\ &= \sum_{S \subseteq N} \sum_{i \in S} \gamma_S w(S) - \sum_{S \subseteq N} \sum_{i \in S} \gamma_S w(S \setminus \{i\}) \\ &= \sum_{S \subseteq N} s \times \gamma_S w(S) - \sum_{S \subseteq N} \gamma_S \left(\sum_{i \in S} w(S \setminus \{i\}) \right) \\ &= \sum_{S \subseteq N} s \times \gamma_S w(S) - \sum_{S \subset N} s \times \gamma_S w(S) \\ &= n \times \gamma_N w(N) \\ &= w(N) \end{aligned}$$

3.9 If $i, j \in S$

$$w(S \setminus \{i\}) = w(S \setminus \{i,j\} \cup \{j\}) = w(S \setminus \{i,j\} \cup \{i\}) = w(S \setminus \{i\})$$

$$\begin{aligned} \varphi_i(w) &= \sum_{S \ni i} \gamma_S (w(S) - w(S \setminus \{i\})) \\ &= \sum_{S \ni i,j} \gamma_S (w(S) - w(S \setminus \{i\})) + \sum_{S \ni i, S \not\ni j} \gamma_S (w(S) - w(S \setminus \{i\})) \\ &= \sum_{S \ni i,j} \gamma_S (w(S) - w(S \setminus \{j\})) + \sum_{S \not\ni i,j} \gamma_S (w(S \cup \{i\}) - w(S)) \\ &= \sum_{S \ni i,j} \gamma_S (w(S) - w(S \setminus \{j\})) + \sum_{S' \not\ni i,j} \gamma_{S'} (w(S' \cup \{j\}) - w(S')) \\ &= \sum_{S \ni i,j} \gamma_S (w(S) - w(S \setminus \{j\})) + \sum_{S \not\ni i, S \ni j} \gamma_S (w(S) - w(S \setminus \{j\})) \\ &= \varphi_j(w) \end{aligned}$$

3.10 For any null player

$$w(S) - w(S \setminus \{i\}) = 0$$

for every $S \subseteq N$. Consequently

$$\varphi_i(w) = \sum_{S \subseteq N} \gamma_S (w(S) - w(S \setminus \{i\})) = 0$$

3.11 Every $i \notin T$ is a null player, so that

$$\varphi_i(u_T) = 0 \qquad \text{for every} \quad i \notin T$$

Feasibility requires that

$$\sum_{i \in T} \varphi_i(u_T) = \sum_{i \in N} \varphi_i(u_T) = 1$$

Further, any two players in T are substitutes, so that symmetry requires that

$$\varphi_i(u_T) = \varphi_j(u_T) \qquad \text{for every} \quad i, j \in T$$

Together, these conditions require that

$$\varphi_i(u_T) = \frac{1}{t} \qquad \text{for every} \quad i \in T$$

The Shapley value of the a T-unanimity game is

$$\varphi_i(u_T) = \begin{cases} \frac{1}{t} & i \in T \\ 0 & i \notin T \end{cases}$$

where $t = |T|$.

3.12 Any coalitional game can be represented as a linear combination of unanimity games u_T (Example 1.75)

$$w = \sum_T \alpha_T u_T$$

By linearity, the Shapley value is

$$\begin{aligned} \varphi w &= \varphi \left(\sum_{T \subseteq N} \alpha_T u_T \right) \\ &= \sum_{T \subseteq N} \alpha_T \varphi u_T \end{aligned}$$

and therefore for player i

$$\begin{aligned} \varphi_i w &= \sum_{T \subseteq N} \alpha_T \varphi_i u_T \\ &= \sum_{\substack{T \subseteq N \\ T \ni i}} \frac{1}{t} \alpha_T \\ &= \sum_{T \subseteq N} \frac{1}{t} \alpha_T - \sum_{\substack{T \subseteq N \\ i \notin T}} \frac{1}{t} \alpha_T \\ &= P(N, w) - P(N \setminus \{i\}, w) \end{aligned}$$

Using Exercise 3.8

$$\begin{aligned} w(N) &= \sum_{i \in N} \varphi_i w \\ &= \sum_{i \in N} \Big(P(N, w) - P(N \setminus \{i\}, v) \Big) \\ &= nP(N, w) - \sum_{i \in N} P(N \setminus \{i\}, v) \end{aligned}$$

which implies that

$$P(N, w) = \frac{1}{n} \left(w(N) - \sum_{i \in N} P(N \setminus \{i\}, v) \right)$$

3.13 Choose any $\mathbf{x} \neq \mathbf{0} \in X$.

$$\mathbf{0}_X = \mathbf{x} - \mathbf{x}$$

and by additivity

$$\begin{aligned} f(\mathbf{0}_X) &= f(\mathbf{x} - \mathbf{x}) \\ &= f(\mathbf{x}) - f(\mathbf{x}) \\ &= \mathbf{0}_Y \end{aligned}$$

3.14 Let $\mathbf{x}^1$, $\mathbf{x}^2$ belong to X. Then

$$\begin{aligned} g \circ f(\mathbf{x}^1 + \mathbf{x}^2) &= g \circ f(\mathbf{x}^1 + \mathbf{x}^2) \\ &= g\left(f(\mathbf{x}^1) + f(\mathbf{x}^2)\right) \\ &= g\left(f(\mathbf{x}^1)\right) + g\left(f(\mathbf{x}^2)\right) \\ &= g \circ f(\mathbf{x}^1) + g \circ f(\mathbf{x}^2) \end{aligned}$$

and

$$\begin{aligned} g \circ f(\alpha \mathbf{x}) &= g\left(f(\alpha \mathbf{x})\right) \\ &= g\left(\alpha f(\mathbf{x})\right) \\ &= \alpha g\left(f(\mathbf{x})\right) \\ &= \alpha g \circ f(\mathbf{x}) \end{aligned}$$

Therefore $g \circ f$ is linear.

3.15 Let S be a subspace of X and let $\mathbf{y}_1$, $\mathbf{y}_2$ belong to $f(S)$. Choose any $\mathbf{x}_1 \in f^{-1}(\mathbf{y}_1)$ and $\mathbf{x}_2 \in f^{-1}(\mathbf{y}_2)$. Then for $\alpha_1, \alpha_2 \in \Re$

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 \in S$$

Since f is linear (Exercise 3.1)

$$\alpha_1 \mathbf{y}_1 + \alpha_2 \mathbf{y}_2 = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2) = f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) \in f(S)$$

$f(S)$ is a subspace.

Let T be a subspace of Y and let $\mathbf{x}_1$, $\mathbf{x}_2$ belong to $f^{-1}(T)$. Let $\mathbf{y}_1 = f(\mathbf{x}_1)$ and $\mathbf{y}_2 = f(\mathbf{x}_2)$. Then $\mathbf{y}_1, \mathbf{y}_2 \in T$. For every $\alpha_1, \alpha_2 \in \Re$

$$\alpha_1 \mathbf{y}_1 + \alpha_2 \mathbf{y}_2 = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2) \in T$$

Since f is linear, this implies that

$$f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2) \in T$$

Therefore

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 \in f^{-1}(T)$$

We conclude that $f^{-1}(T)$ is a subspace.

3.16 $f(X)$ is a subspace of Y. rank $f(X) =$ rank Y implies that $f(X) = Y$. f is onto.

3.17 This is a special case of the previous exercise, since $\{\mathbf{0}_Y\}$ is a subspace of Y.

3.18 Assume not. That is, assume that there exist two distinct elements $\mathbf{x}_1$ and $\mathbf{x}_2$ with $f(\mathbf{x}_1) = f(\mathbf{x}_2)$. Then $\mathbf{x}_1 - \mathbf{x}_2 \neq \mathbf{0}_X$ but

$$f(\mathbf{x}_1 - \mathbf{x}_2) = f(\mathbf{x}_1) - f(\mathbf{x}_2) = \mathbf{0}_Y$$

so that $\mathbf{x}_1 - \mathbf{x}_2 \in$ kernel f which contradicts the assumption that kernel $f = \{\mathbf{0}\}$.

3.19 If f has an inverse, then it is one-to-one and onto (Exercise 2.4), that is $f^{-1}(\mathbf{0}) = \mathbf{0}$ and $f(X) = Y$. Conversely, if kernel $f = \{\mathbf{0}\}$ then f is one-to-one by the previous exercise. If furthermore $f(X) = Y$, then f is one-to-one and onto, and therefore has an inverse (Exercise 2.4).

3.20 Let f be a nonsingular linear function from X to Y with inverse f^{-1}. Choose $\mathbf{y}^1, \mathbf{y}^2 \in Y$ and let

$$\begin{aligned} \mathbf{x}_1 &= f^{-1}(\mathbf{y}_1) \\ \mathbf{x}_2 &= f^{-1}(\mathbf{y}_2) \end{aligned}$$

so that

$$\begin{aligned} \mathbf{y}_1 &= f(\mathbf{x}_1) \\ \mathbf{y}_2 &= f(\mathbf{x}_2) \end{aligned}$$

Since f is linear

$$f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2) = \mathbf{y}_1 + \mathbf{y}_2$$

which implies that

$$f^{-1}(\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{x}_1 + \mathbf{x}_2 = f^{-1}(\mathbf{y}_1) + f^{-1}(\mathbf{y}_2)$$

The homogeneity of f^{-1} can be demonstrated similarly.

3.21 Assume that $f \colon X \to Y$ and $g \colon Y \to Z$ are nonsingular. Then (Exercise 3.19)

- $f(X) = Y$ and $g(Y) = Z$
- kernel $f = \{0_X\}$ and kernel $g = \{0_Y\}$

We have previously shown (Exercise 3.14) that $h = g \circ f \colon X \to Z$ is linear. To show that h is nonsingular, we note that

- $h(X) = g \circ f(X) = g(Y) = Z$

- If $\mathbf{x} \in \text{kernel}\,(h)$ then

$$h(\mathbf{x}) = g\left(f(\mathbf{x})\right) = \mathbf{0}$$

and $f(\mathbf{x}) \in \text{kernel}\, g = \{\mathbf{0}_Y\}$. Therefore $f(\mathbf{x}) = \mathbf{0}_Y$ which implies that $\mathbf{x} = \mathbf{0}_X$. Thus kernel $h = \{\mathbf{0}_X\}$.

We conclude that h is nonsingular.

Finally, let $\mathbf{z}$ be any point in Z and let

$$\begin{aligned} \mathbf{x}_1 &= h^{-1}(\mathbf{z}) = (g \circ f)^{-1}(\mathbf{z}) \\ \mathbf{y} &= g^{-1}(\mathbf{z}) \\ \mathbf{x}_2 &= f^{-1}(\mathbf{y}) \end{aligned}$$

Then

$$\begin{aligned} \mathbf{z} &= h(\mathbf{x}_1) = g \circ f(\mathbf{x}_1) \\ \mathbf{z} &= g(\mathbf{y}) = g \circ f(\mathbf{x}_2) \end{aligned}$$

which implies that $\mathbf{x}_1 = \mathbf{x}_2$.

3.22 Suppose f were one-to-one. Then kernel $f = \{\mathbf{0}\} \subseteq$ kernel h and $g = h \circ f^{-1}$ is a well-defined linear function mapping $f(X)$ to Y with

$$g \circ f = \left(h \circ f^{-1}\right) \circ f = h$$

We need to show that this still holds if f is *not* one-to-one. In this case, for arbitrary $\mathbf{y} \in f(X)$, $f^{-1}(\mathbf{y})$ may contain more than one element. Suppose $\mathbf{x}_1$ and $\mathbf{x}_2$ are distinct elements in $f^{-1}(\mathbf{y})$. Then

$$f(\mathbf{x}_1 - \mathbf{x}_2) = f(\mathbf{x}_1) - f(\mathbf{x}_2) = \mathbf{y} - \mathbf{y} = \mathbf{0}$$

so that $\mathbf{x}_1 - \mathbf{x}_2 \in$ kernel $f \subseteq$ kernel h (by assumption). Therefore

$$h(\mathbf{x}_1) - h(\mathbf{x}_2) = h(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$$

which implies that $h(\mathbf{x}_1) = h(\mathbf{x}_2)$ for all $\mathbf{x}_1, \mathbf{x}_2 \in f^{-1}(\mathbf{y})$. Thus $g = h \circ f^{-1} \colon f(X) \to Z$ is well defined even if f is many-to-one.

To show that g is linear, choose $\mathbf{y}_1$, $\mathbf{y}_2$ in $f(X)$ and let

$$\begin{aligned} \mathbf{x}_1 &\in f^{-1}(\mathbf{y}_1) \\ \mathbf{x}_2 &\in f^{-1}(\mathbf{y}_2) \end{aligned}$$

Since $f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2) = \mathbf{y}_1 + \mathbf{y}_2$

$$\mathbf{x}_1 + \mathbf{x}_2 \in f^{-1}(\mathbf{y}_1 + \mathbf{y}_2)$$

and

$$g(\mathbf{y}_1 + \mathbf{y}_2) = h(\mathbf{x}_1 + \mathbf{x}_2)$$

Therefore

$$\begin{aligned} g(\mathbf{y}_1) + g(\mathbf{y}_2) &= h(\mathbf{x}_1) + h(\mathbf{x}_2) \\ &= h(\mathbf{x}_1 + \mathbf{x}_2) \\ &= g(\mathbf{y}_1 + \mathbf{y}_2) \end{aligned}$$

Similarly $\alpha \mathbf{x}_1 \in f^{-1}(\alpha \mathbf{y}_1)$ and

$$\begin{aligned} g(\alpha \mathbf{y}_1) &= h(\alpha \mathbf{x}_1) \\ &= \alpha h(\mathbf{x}_1) \\ &= \alpha g(\mathbf{y}_1) \end{aligned}$$

We conclude that $g = h \circ f^{-1}$ is a linear function mapping $f(X)$ to Z with $h = g \circ f$.

3.23 Let $\mathbf{y}$ be an arbitrary element of $f(X)$ with $\mathbf{x} \in f^{-1}(\mathbf{y})$. Since B is a basis for X, $\mathbf{x}$ can be represented as a linear combination of elements of B, that is there exists $\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_m \in B$ and $\alpha_1, ..., \alpha_m \in R$ such that

$$\begin{aligned} \mathbf{x} &= \sum_{i=1}^{m} \alpha_i \mathbf{x}_i \\ \mathbf{y} &= f(\mathbf{x}) \\ &= f\left(\sum_i \alpha_i \mathbf{x}_i\right) \\ &= \sum_i \alpha_i f(\mathbf{x}_i) \end{aligned}$$

Since $f(\mathbf{x}_i) \in f(B)$, we have shown that $\mathbf{y}$ can be written as a linear combination of elements of $f(B)$, that is

$$\mathbf{y} \in \text{lin } B$$

Since the choice of $\mathbf{y}$ was arbitrary, $f(B)$ spans $f(X)$, that is

$$\text{lin } B = f(X)$$

3.24 Let $n = \dim X$ and $k = \dim \text{kernel } f$. Let $\mathbf{x}_1, \dots, \mathbf{x}_k$ be a basis for the kernel of f. This can be extended (Exercise 1.142) to a basis B for X. Exercise 3.23 showed

$$\text{lin } B = f(X)$$

Since $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \text{kernel } f$, $f(\mathbf{x}_i) = \mathbf{0}$ for $i = 1, 2, \dots, k$. This implies that $\{f(\mathbf{x}_{k+1}), \dots, f(\mathbf{x}_n)\}$ spans $f(X)$, that is

$$\text{lin } \{(\mathbf{x}_{k+1}), ..., f(\mathbf{x}_n)\} = f(X)$$

To show that $\dim f(X) = n - k$, we have to show that $\{f(x_{k+1}), f(x_{k+2}), \dots, f(x_n)\}$ is linearly independent. Assume not. That is, assume there exist $\alpha_{k+1}, \alpha_{k+2}, ..., \alpha_n \in R$ such that

$$\sum_{i=k+1}^{n} \alpha_i f(\mathbf{x}_i) = \mathbf{0}$$

This implies that

$$f\left(\sum_{i=k+1}^{n} \alpha_i \mathbf{x}_i\right) = \mathbf{0}$$

or

$$\mathbf{x} = \sum_{i=k+1}^{n} \alpha_i x_i \in \text{kernel } f$$

This implies that $\mathbf{x}$ can also be expressed as a linear combination of elements in $\{\mathbf{x}_1, x_2, ..., \mathbf{x}_k\}$, that is there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_k$ such that

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$$

or

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \sum_{i=k+1}^{n} \alpha_i \mathbf{x}_i$$

which contradicts the assumption that B is a basis for X. Therefore $\{f(\mathbf{x}_{k+1}), \dots, f(\mathbf{x}_n)\}$ is a basis for $f(X)$ and therefore $\dim f(x) = n - k$. We conclude that

$$\dim \operatorname{kernel} f + \dim f(X) = n = \dim X$$

3.25 Equation (3.2) implies that nullity $f = 0$, and therefore f is one-to-one (Exercise 3.18).

3.26 Choose some $\mathbf{x} = (x_1, x_2, \dots, x_n) \in X$. $\mathbf{x}$ has a unique representation in terms of the standard basis (Example 1.79)

$$\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j$$

Let $\mathbf{y} = f(\mathbf{x})$. Since f is linear

$$\mathbf{y} = f(\mathbf{x}) = f\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} \mathbf{x}_j f(\mathbf{e}_j)$$

Each $f(\mathbf{e}_j)$ has a unique representation of the form

$$f(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{e}_i$$

so that

$$\begin{aligned}
\mathbf{y} &= f(\mathbf{x}) \\
&= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} \mathbf{e}_i\right) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) \mathbf{e}_i \\
&= \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix} \\
&= A\mathbf{x}
\end{aligned}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

3.27

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

3.28 We must specify bases for each space. The most convenient basis for G^N is the T-unanimity games. We adopt the standard basis for $\Re^n$. With respect to these bases, the Shapley value φ is represented by the $2^{n-1} \times n$ matrix where each row is the Shapley value of the corresponding T-unanimity game.

For three player games ($n = 3$), the matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

3.29 Clearly, if f is continuous, f is continuous at $\mathbf{0}$.

To show the converse, assume that $f: X \to Y$ is continuous at $\mathbf{0}$. Let $(\mathbf{x}^n)$ be a sequence which converges to $\mathbf{x} \in X$. Then the sequence $(\mathbf{x}^n - \mathbf{x})$ converges to $\mathbf{0}_X$ and therefore $f(\mathbf{x}^n - \mathbf{x}) \to \mathbf{0}_Y$ by continuity (Exercise 2.68). By linearity, $f(\mathbf{x}^n) - f(\mathbf{x}) = f(\mathbf{x}^n - \mathbf{x}) \to \mathbf{0}_Y$ and therefore $f(\mathbf{x}^n)$ converges to $f(\mathbf{x})$. We conclude that f is continuous at $\mathbf{x}$.

3.30 Assume that f is bounded, that is

$$\|f(\mathbf{x})\| \leq M \|\mathbf{x}\| \text{ for every } \mathbf{x} \in X$$

Then f is Lipschitz at $\mathbf{0}$ (with Lipschitz constant M) and hence continuous (by the previous exercise).

Conversely, assume f is continuous but not bounded. Then, for every positive integer n, there exists some $\mathbf{x}^n \in X$ such that $\|f(\mathbf{x}^n)\| > n \|\mathbf{x}^n\|$ which implies that

$$\left\| f\left(\frac{\mathbf{x}^n}{n \|\mathbf{x}_n\|} \right) \right\| > 1$$

Define

$$\mathbf{y}^n = \frac{\mathbf{x}^n}{n \|\mathbf{x}_n\|}$$

Then $\mathbf{y}^n \to \mathbf{0}$ but $f(\mathbf{y}^n) \not\to \mathbf{0}$. This implies that f is not continuous at the origin, contradicting our hypothesis.

3.31 Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for X. For every $\mathbf{x} \in X$, there exists numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$$

and

$$\begin{aligned} f(\mathbf{x}) &= \sum_{i=1}^{n} \alpha_i f(\mathbf{x}_i) \\ \|f(\mathbf{x})\| &= \left\| \sum_{i=1}^{n} \alpha_i f(\mathbf{x}_i) \right\| \\ &\leq \sum_{i=1}^{n} |\alpha_i| \, \|f(\mathbf{x}_i)\| \\ &\leq \left(\max_{i=1}^{n} \|f(\mathbf{x}_i)\| \right) \sum_{i=1}^{n} |\alpha_i| \end{aligned}$$

By Lemma 1.1, there exists a constant c such that

$$\sum_{i=1}^{n} |\alpha_i| \leq \frac{1}{c} \left\| \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \right\| = \frac{1}{c} \|\mathbf{x}\|$$

Combining these two inequalities

$$\|f(\mathbf{x})\| \leq M \|\mathbf{x}\|$$

where $M = \max_{i=1}^{n} \|f(\mathbf{x}_i)\| / c$.

3.32 For any $\mathbf{x} \in X$, let $a = \|\mathbf{x}\|$ and define $\mathbf{y} = \mathbf{x}/a$. Linearity implies that

$$\|f\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|f(\mathbf{x})\|}{a} = \sup_{\mathbf{x} \neq \mathbf{0}} \|f(\mathbf{x}/a)\| = \sup_{\|\mathbf{y}\|=1} \|f(\mathbf{y})\|$$

3.33 **$\|f\|$ is a norm** Let $f \in BL(X, Y)$. Clearly

$$\|f\| = \sup_{\|\mathbf{x}\|=1} \|f(\mathbf{x})\| \geq 0$$

Further, for every $\alpha \in \Re$,

$$\|\alpha f\| = \sup_{\|\mathbf{x}\|=1} \|\alpha f(\mathbf{x})\| = |\alpha| \, \|f\|$$

Finally, for every $g \in BL(X, Y)$,

$$\|f + g\| = \sup_{\|\mathbf{x}\|=1} \|f(\mathbf{x}) + g(\mathbf{x})\| \leq \sup_{\|\mathbf{x}\|=1} \|f(\mathbf{x})\| + \sup_{\|\mathbf{x}\|=1} \|g(\mathbf{x})\| \leq \|f\| + \|g\|$$

verifying the triangle inequality. There $\|f\|$ is a norm.

$BL(X, Y)$ is a linear space Let $f, g \in BL(X, Y)$. Since $BL(X, Y) \subseteq L(X, Y)$, $f + g$ is linear, that is $f + g \in L(X, Y)$ (Exercise 3.2). Similarly, $\alpha f \in L(X, Y)$ for every $\alpha \in \Re$. Further, by the triangle inequality $\|f + g\| \leq \|f\| + \|g\|$ and therefore for every $\mathbf{x} \in X$

$$\|(f + g)(\mathbf{x})\| \leq \|f + g\| \, \|\mathbf{x}\| \leq (\|f\| + \|g\|) \, \|\mathbf{x}\|$$

Therefore $f + g \in BL(X, Y)$. Similarly

$$\|(\alpha f)(\mathbf{x})\| \leq (|\alpha| \, \|f\|) \, \|\mathbf{x}\|$$

so that $\alpha f \in BL(X, Y)$.

$BL(X,Y)$ is complete with this norm Let (f^n) be a Cauchy sequence in $BL(X,Y)$. For every $\mathbf{x} \in X$

$$\|f^n(\mathbf{x}) - f^m(\mathbf{x})\| \leq \|f^n - f^m\| \, \|\mathbf{x}\|$$

Therefore $(f^n(\mathbf{x}))$ is a Cauchy sequence in Y, which converges since Y is complete. Define the function $f \colon X \to Y$ by $f(\mathbf{x}) = \lim_{n\to\infty} f^n(\mathbf{x})$.

f is linear since

$$f(\mathbf{x}_1 + \mathbf{x}_2) = \lim f^n(\mathbf{x}_1 + \mathbf{x}_2) = \lim f^n(\mathbf{x}_1) + \lim f^n(\mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2)$$

and

$$f(\alpha\mathbf{x}) = \lim f^n(\alpha\mathbf{x}) = \alpha \lim f^n(\mathbf{x}) = \alpha f(\mathbf{x})$$

To show that f is bounded, we observe that

$$\|f(\mathbf{x})\| = \left\| \lim_n f^n(\mathbf{x}) \right\| = \lim_n \|f^n(\mathbf{x})\| \leq \sup_n \|f^n(\mathbf{x})\| \leq \sup_n \|f^n\| \, \|\mathbf{x}\|$$

Since (f^n) is a Cauchy sequence, (f^n) is bounded (Exercise 1.100), that is there exists M such that $\|f^n\| \leq M$. This implies

$$\|f(\mathbf{x})\| \leq \sup_n \|f^n\| \, \|\mathbf{x}\| \leq M \, \|\mathbf{x}\|$$

Thus, f is bounded.

To complete the proof, we must show $f^n \to f$, that is $\|f^n - f\| \to 0$. Since (f^n) is a Cauchy sequence, for every $\epsilon > 0$, there exists N such that $\|f^n - f^m\| \leq \epsilon$ for every $n, m \geq N$ and consequently

$$\|f^n(\mathbf{x}) - f^m(\mathbf{x})\| = \|(f^n - f^m)(\mathbf{x})\| \leq \epsilon \, \|\mathbf{x}\|$$

Letting m go to infinity,

$$\|f^n(\mathbf{x}) - f(\mathbf{x})\| = \|(f^n - f)(\mathbf{x})\| \leq \epsilon \, \|\mathbf{x}\|$$

for every $\mathbf{x} \in X$ and $n \geq N$ and therefore

$$\|f^n - f\| = \sup_{\|\mathbf{x}\|=1} \{f^n - f)(\mathbf{x})\} \leq \epsilon$$

for every $n \geq N$.

3.34 1. Since X is finite-dimensional, S is compact (Proposition 1.4). Since f is continuous, $f(S)$ is a compact set in Y (Exercise 2.3). Since $\mathbf{0}_X \notin S$, $\mathbf{0}_Y = f(\mathbf{0}_X) \notin f(S)$.

2. Consequently, $\big(f(S)\big)^c$ is an open set containing $\mathbf{0}_Y$. It contains an open ball $T \subseteq \big(f(S)\big)^c$ around $\mathbf{0}_Y$.

3. Let $\mathbf{y} \in T$ and choose any $\mathbf{x} \in f^{-1}(\mathbf{y})$ and consider $\mathbf{y}/\|\mathbf{x}\|$. Since f is linear,

$$\frac{\mathbf{y}}{\|\mathbf{x}\|} = \frac{f(\mathbf{x})}{\|\mathbf{x}\|} = f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) \in f(S)$$

and therefore $\mathbf{y}/\|\mathbf{x}\| \notin T$ since $T \cap f(S) = \emptyset$.

Suppose that $\mathbf{y} \notin f(B)$. Then $\|\mathbf{x}\| \geq 1$ and therefore

$$\mathbf{y} \in T \implies \frac{\mathbf{y}}{\|\mathbf{x}\|} \in T$$

since T is convex. This contradiction establishes that $\mathbf{y} \in f(B)$ and therefore $T \subseteq f(B)$. We conclude that $f(B)$ contains an open ball around $\mathbf{0}_Y$.

4. Let S be any open set in X. We need to show that $f(S)$ is open in Y. Choose any $\mathbf{y} \in f(S)$ and $\mathbf{x} \in f^{-1}(\mathbf{y})$. Then $\mathbf{x} \in S$ and, since S is open, there exists some $r > 0$ such that $B_r(\mathbf{x}) \subseteq S$. Now $B_r(\mathbf{x}) = \mathbf{x} + rB$ and

$$f(B_r(\mathbf{x})) = \mathbf{y} + rf(B) \subseteq f(S)$$

by linearity. As we have just shown, there exists an open ball T about $\mathbf{0}_Y$ such that $T \subseteq f(B)$. Let $T(\mathbf{x}) = \mathbf{y} + rT$. $T(\mathbf{x})$ is an open ball about $\mathbf{y}$. Since $T \subseteq f(B)$, $T(\mathbf{x}) = \mathbf{y} + rT \subseteq f(B_r(\mathbf{x})) \subseteq f(S)$. This implies that $f(S)$ is open. Since S was an arbitrary open set, f is an open map.

5. Exercise 2.69.

3.35 ***f* is linear**

$$f(\alpha + \beta) = \sum_{i=1}^{n} (\alpha_i + \beta_i)\mathbf{x}_i = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i + \sum_{i=1}^{n} \beta_i \mathbf{x}_i = f(\alpha) + f(\beta)$$

Similarly for every $t \in \Re$

$$f(t\alpha) = t\sum_{i=1}^{n} \alpha_i \mathbf{x}_i = tf(t\alpha)$$

***f* is one-to-one** Exercise 1.137.

***f* is onto** By definition of a basis

$$\text{lin}\,\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} = X$$

***f* is continuous** Exercise 3.31

***f* is an open map** Proposition 3.2

3.36 f is bounded and therefore there exists M such that $\|f(\mathbf{x})\| \leq M\|\mathbf{x}\|$. Similarly, f^{-1} is bounded and therefore there exists m such that for every $\mathbf{x}$

$$f^{-1}(\mathbf{y}) \leq \frac{1}{m}\|\mathbf{y}\|$$

where $\mathbf{y} = f(\mathbf{x})$. This implies

$$m\|\mathbf{x}\| \leq \|f(\mathbf{x})\|$$

and therefore for every $\mathbf{x} \in X$.

$$m\|\mathbf{x}\| \leq \|f(\mathbf{x})\| \leq M\|\mathbf{x}\|$$

By the linearity of f,

$$m\|\mathbf{x}_1 - \mathbf{x}_2\| \leq \|f(\mathbf{x}_1 - \mathbf{x}_2)\| = \|f(\mathbf{x}_1) - f(\mathbf{x}_2)\| \leq M\|\mathbf{x}_1 - \mathbf{x}_2\|$$

3.37 For any function, continuity implies closed graph (Exercise 2.70). To show the converse, assume that $G = \text{graph}(f)$ is closed. $X \times Y$ with norm $\|(\mathbf{x}, \mathbf{y})\| = \max\{\|\mathbf{x}\|, \|\mathbf{y}\|\}$ is a Banach space (Exercise 1.209). Since G is closed, G is complete. Also, G is a subspace of $X \times Y$. Consequently, G is a Banach space in its own right.

Consider the projection $h\colon G \to X$ defined by $h(\mathbf{x}, f(\mathbf{x})) = \mathbf{x}$. Clearly h is linear, one-to-one and onto with

$$h^{-1}(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$$

It is also bounded since

$$\|h(\mathbf{x}, f(\mathbf{x}))\| = \|\mathbf{x}\| \le \|(\mathbf{x}, f(\mathbf{x})\|$$

By the Open Mapping Theorem, h^{-1} is bounded. For every $\mathbf{x} \in X$

$$\|f(\mathbf{x})\| \le \|(\mathbf{x}, f(\mathbf{x}))\| = \left\|h^{-1}(\mathbf{x})\right\| \le \left\|h^{-1}\right\| \|\mathbf{x}\|$$

We conclude that f is bounded and hence continuous.

3.38 $f(1) = 5$, $f(2) = 7$ but

$$f(1+2) = f(3) = 9 \neq f(1) + f(2)$$

Similarly

$$f(3 \times 2) = f(6) = 15 \neq 3 \times f(2)$$

3.39 Assume f is affine. Let $\mathbf{y} = f(\mathbf{0})$ and define

$$g(\mathbf{x}) = f(\mathbf{x}) - \mathbf{y}$$

g is homogeneous since for every $\alpha \in \Re$

$$\begin{aligned}
g(\alpha\mathbf{x}) &= g(\alpha\mathbf{x} + (1-\alpha)\mathbf{0}) \\
&= f(\alpha\mathbf{x} + (1-\alpha)\mathbf{0}) - \mathbf{y} \\
&= \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{0}) - \mathbf{y} \\
&= \alpha f(\mathbf{x}) + (1-\alpha)\mathbf{y} - \mathbf{y} \\
&= \alpha f(\mathbf{x}) - \alpha\mathbf{y} \\
&= \alpha(f(x) - \mathbf{y}) \\
&= \alpha g(\mathbf{x})
\end{aligned}$$

Similarly for any $\mathbf{x}_1, \mathbf{x}_2 \in X$

$$\begin{aligned}
g(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &= f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) - y \\
&= \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) - y
\end{aligned}$$

Therefore, for $\alpha = 1/2$

$$\begin{aligned}
g(\frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2) &= \frac{1}{2} f(\mathbf{x}_1) + \frac{1}{2} f(\mathbf{x}_2) - y \\
&= \frac{1}{2}(f(\mathbf{x}_1) - y) + \frac{1}{2}(f(\mathbf{x}_2) - y) \\
&= \frac{1}{2} g(\mathbf{x}_1) + \frac{1}{2} g(\mathbf{x}_2)
\end{aligned}$$

Since g is homogeneous

$$g(\mathbf{x}_1 + \mathbf{x}_2) = g(\mathbf{x}_1) + g(\mathbf{x}_2)$$

which shows that g is additive and hence linear.

Conversely if

$$f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{y}$$

with g linear

$$\begin{aligned} f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &= \alpha g(\mathbf{x}_1) + (1-\alpha) g(\mathbf{x}_2) + y \\ &= \alpha g(\mathbf{x}_1) + y + (1-\alpha) g(\mathbf{x}_2) + y \\ &= \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \end{aligned}$$

3.40 Let S be an affine subset of X and let $\mathbf{y}_1$, $\mathbf{y}_2$ belong to $f(S)$. Choose any $\mathbf{x}_1 \in f^{-1}(\mathbf{y}_1)$ and $\mathbf{x}_2 \in f^{-1}(\mathbf{y}_2)$. Then for any $\alpha \in \Re$

$$\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in S$$

Since f is affine

$$\alpha \mathbf{y}_1 + (1-\alpha)\mathbf{y}_2 = \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) = f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \in f(S)$$

$f(S)$ is an affine set.

Let T be an affine subset of Y and let $\mathbf{x}_1$, $\mathbf{x}_2$ belong to $f^{-1}(T)$. Let $\mathbf{y}_1 = f(\mathbf{x}_1)$ and $\mathbf{y}_2 = f(\mathbf{x}_2)$. Then $\mathbf{y}_1, \mathbf{y}_2 \in T$. For every $\alpha \in \Re$

$$\alpha \mathbf{y}_1 + (1-\alpha)\mathbf{y}_2 = \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \in T$$

Since f is affine, this implies that

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) = \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \in T$$

Therefore

$$\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in f^{-1}(T)$$

We conclude that $f^{-1}(T)$ is an affine set.

3.41 For any $\mathbf{y}_1, \mathbf{y}_2 \in f(S)$, choose $\mathbf{x}_1, \mathbf{x}_2 \in S$ such that $\mathbf{y}_i = f(\mathbf{x}_i)$. Since S is convex, $\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in S$ and therefore

$$\begin{aligned} f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &= \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \\ &= \alpha \mathbf{y}_1 + (1-\alpha)\mathbf{y}_2 \in f(S) \end{aligned}$$

Therefore $f(S)$ is convex.

3.42 Suppose otherwise that $\mathbf{y}$ is not efficient. Then there exists another production plan $\mathbf{y}' \in Y$ such that $\mathbf{y}' \geq \mathbf{y}$. Since $\mathbf{p} > 0$, this implies that $\mathbf{p}\mathbf{y}' > \mathbf{p}\mathbf{y}$, contradicting the assumption that $\mathbf{y}$ maximizes profit.

3.43 The random variable X can be represented as the sum

$$X = \sum_{s \in S} X(s) \chi_{\{s\}}$$

where $\chi_{\{s\}}$ is the indicator function of the set $\{s\}$. Since E is linear

$$E(X) = \sum_{s \in S} X(s) E(\chi_{\{s\}})$$
$$= \sum_{s \in S} p_S X(s)$$

since $E(\chi_{\{s\}} = P(\{s\}) = p_s \geq 0$. For the random variable $X = \mathbf{1}$, $X(s) = 1$ for every $s \in S$ and

$$E(\mathbf{1}) = \sum_{s \in S} p_S = 1$$

3.44 Let $x_1, x_2 \in C[0,1]$. Recall that addition in C[0,1] is defined by

$$(x_1 + x_2)(t) = x_1(t) + x_2(t)$$

Therefore

$$f(x_1 + x_2) = (x_1 + x_2)(1/2) = x_1(1/2) + x_2(1/2) = f(x_1) + f(x_2)$$

Similarly

$$f(\alpha x_1) = (\alpha x_1)(1/2) = \alpha x_1(1/2) = \alpha f(x_1)$$

3.45 Assume that $\mathbf{x}^* = \mathbf{x}_1^* + \mathbf{x}_2^* + \cdots + \mathbf{x}_n^*$ maximizes f over S. Suppose to the contrary that there exists $\mathbf{y}_j \in S_j$ such that $f(\mathbf{y}_j) > f(\mathbf{x}_j^*)$. Then $\mathbf{y} = \mathbf{x}_1^* + \mathbf{x}_2^* + \cdots + \mathbf{y}_j + \cdots + \mathbf{x}_n^* \in S$ and

$$f(\mathbf{y}) = \sum_{i \neq j} f(\mathbf{x}_i^*) + f(\mathbf{y}_i) > \sum_i f(\mathbf{x}_i^*) = f(\mathbf{x}^*)$$

contradicting the assumption at f is maximized at $\mathbf{x}^*$.

Conversely, assume

$$f(\mathbf{x}_i^*) \geq f(\mathbf{x}_i) \text{ for every } \mathbf{x}_i \in S_i$$

for every $i = 1, 2, \ldots, n$. Summing

$$f(\mathbf{x}^*) = f(\sum \mathbf{x}_i^*) = \sum f(x_i^*) \geq \sum f(\mathbf{x}_i) = f(\sum \mathbf{x}_i) = f(\mathbf{x}) \text{ for every } \mathbf{x} \in S$$

$\mathbf{x}^* = \mathbf{x}_1^* + \mathbf{x}_2^* + \cdots + \mathbf{x}_n^*$ maximizes f over S.

3.46 1. Assume (x_t) is a sequence in l_1 with $s = \sum_{t=1}^{\infty} |x_j| < \infty$. Let (s_t) denote the sequence of partial sums

$$s_t = \sum_{j=1}^{t} |x_j|$$

Then (s_t) is a bounded monotone sequence in $\Re^n$ which converges to s. Consequently, (s_t) is a Cauchy sequence. For every $\epsilon > 0$ there exists an N such that

$$\sum_{n=m}^{m+k} |x_t| < \epsilon$$

for every $m \geq N$ and $k \geq 0$. Letting $k = 0$

$$|x_t| < \epsilon \text{ for every } \; n \geq N$$

We conclude that $x_t \to 0$ so that $(x_t) \in c_0$. This establishes $l_1 \subseteq c_0$.

To see that the inclusion is strict, that is $l_1 \subset c_0$, we observe that the sequence $(1/n) = (1, 1/2, 1/3, \dots)$ converges to zero but that since

$$\sum_{n=1}^{\infty} \left| \frac{1}{n} \right| = 1 + \frac{1}{2} + \frac{1}{3} + = \infty$$

$(1/n) \notin l_1$.

Every convergent sequence is bounded (Exercise 1.97). Therefore $c_0 \subset l_\infty$.

2. Clearly, every sequence $(p_t) \in l_1$ defines a linear functional $f \in c_0'$ given by

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} p_t x_t$$

for every $\mathbf{x} = (x_t) \in c_0$. To show that f is bounded we observe that every $(x_t) \in c_0$ is bounded and consequently

$$|f(\mathbf{x})| \leq \sum_{n=1}^{\infty} |p_t| \, |x_t| \leq \|(x_t)\|_\infty \sum_{n=1}^{\infty} |p_t| = \|(p_t)\|_1 \, \|(x_t)\|_\infty$$

Therefore $f \in c_0^*$.

To show the converse, let $\mathbf{e}_t$ denote the unit sequences

$$\begin{aligned} \mathbf{e}^1 &= (1, 0, 0, \dots) \\ \mathbf{e}^2 &= (0, 1, 0, \dots) \\ \mathbf{e}^3 &= (0, 0, 1, \dots) \end{aligned}$$

$\{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3, \dots, \}$ form a basis for c_0. Then every sequence $(x_t) \in c_0$ has a unique representation

$$(x_t) = \sum_{n=1}^{\infty} x_t \mathbf{e}_t$$

Let $f \in c_0^*$ be a continuous linear functional on c_0. By continuity and linearity

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} x_t f(\mathbf{e}_t)$$

Let

$$p_t = f(\mathbf{e}_t)$$

so that

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} p_t x_t$$

(Every linear function is determined by its action on a basis (Exercise 3.23))

We need to show that the sequence $(p_t) \in l_1$. For any N, consider the sequence $\mathbf{x}_t = (x^1, x^2, \dots, x_t, 0, 0, \dots)$ where

$$x_t = \begin{cases} 0 & p_t = 0 \text{ or } n \geq N \\ \dfrac{|p_t|}{p_t} & \text{otherwise} \end{cases}$$

Then $(\mathbf{x}_t) \in c_0$, $\|\mathbf{x}_t\|_\infty = 1$ and

$$f(\mathbf{x}_t) = \sum_{n=1}^{t} p_t x_t = \sum_{n=1}^{t} |p_t|$$

Since $f \in c_0^*$, f is bounded and therefore

$$f(\mathbf{x}_t) \leq \|f\| \|\mathbf{x}_t\| = \|f\| < \infty$$

and therefore

$$\sum_{n=1}^{t} |p_t| < \infty \text{ for every } \; N = 1, 2, \dots$$

Consequently

$$\sum_{n=1}^{\infty} |p_t| = \sup_N \sum_{n=1}^{t} |p_t| \leq \|f\| < \infty$$

We conclude that $(p_t) \in l_1$ and therefore $c_0^* = l_1$

3. Similarly, every sequence $(p_t) \in l_\infty$ defines a linear functional f on l_1 given by

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} p_t x_t$$

for every $\mathbf{x} = (x_t) \in l_1$. Moreover f is bounded since

$$|f(\mathbf{x})| \leq \sum_{n=1}^{\infty} |p_t| |x_t| \leq \|(p_t)\| \sum_{n=1}^{\infty} |x_t| < \infty$$

for every $\mathbf{x} = (x_t) \in l_1$ Again, given any linear functional $f \in l_1^*$, let $p_t = f(\mathbf{e}_t)$ where $\mathbf{e}_t$ is the n unit sequence. Then f has the representation

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} p_t x_t$$

To show that $(p_t) \in l_\infty$, for $N = 1, 2, \dots$, consider the sequence $\mathbf{x}_t = (0, 0, \dots, x_t, 0, 0, \dots)$ where

$$x_t = \begin{cases} \dfrac{|p_t|}{p_t} & n = N \text{ and } \; p_t \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then $\mathbf{x}_t \in l_1$, $\|\mathbf{x}_t\|_1 = 1$ and

$$f(\mathbf{x}_t) = |p_t|$$

Since $f \in l_1^*$, f is bounded and therefore

$$|p^N| = f(\mathbf{x}^N) \leq \|f\| \|\mathbf{x}^n\| = \|f\|$$

for every N. Consequently $(p^N) \in l_\infty$. We conclude that $l_1^* = l_\infty$

3.47 By linearity

$$\begin{aligned}\varphi(x,t) &= \varphi(x,0) + \varphi(0,t)\\ &= \varphi(x,0) + \varphi(0,1)t\end{aligned}$$

Considered as a function of x, $\varphi(x,0)$ is a linear functional on X. Define

$$\begin{aligned}g(x) &= \varphi(x,0)\\ \alpha &= \varphi(0,1)\end{aligned}$$

Then

$$\varphi(x,t) = g(x) + \alpha t$$

3.48 Suppose

$$\bigcap_{j=1}^{m} \text{kernel } g_j \subseteq \text{kernel } f$$

Define the function $G \colon X \to \Re^n$ by

$$G(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x}))$$

Then

$$\begin{aligned}\text{kernel } G &= \{\, \mathbf{x} \in X : g_j(\mathbf{x}) = 0, j = 1, 2, \dots m \,\}\\ &= \bigcap_{j=1}^{m} \text{kernel } g_j\\ &\subseteq \text{kernel } f\end{aligned}$$

$f \colon X \to \Re$ and $G \colon X \to \Re^n$. By Exercise 3.22, there exists a linear function $H \colon \Re^n \to \Re$ such that $f = H \circ G$. That is, for every $x \in X$

$$f(\mathbf{x}) = H \circ G(\mathbf{x}) = H(g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x}))$$

Let $\alpha_j = H(\mathbf{e}_j)$ where $\mathbf{e}_j$ is the j-th unit vector in $\Re^m$. Since every linear mapping is determined by its action on a basis, we must have

$$f(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \dots + \alpha_m g_m(\mathbf{x}) \qquad \text{for every } \; x \in X$$

That is

$$f \in \text{lin } g_1, g_2, \dots, g_m$$

Conversely, suppose

$$f \in \text{lin } g_1, g_2, \dots, g_m$$

That is

$$f(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \dots + \alpha_m g_m(\mathbf{x}) \qquad \text{for every } \; x \in X$$

For every $\mathbf{x} \in \bigcap_{j=1}^{m} \text{kernel } g_j$, $g_j(\mathbf{x}) = 0$, $j = 1, 2, \dots, m$ and therefore $f(\mathbf{x}) = 0$. Therefore $x \in \text{kernel } f$. That is

$$\bigcap_{j=1}^{m} \text{kernel } g_j \subseteq \text{kernel } f$$

3.49 Let H be a hyperplane in X. Then there exists a unique subspace V such that $H = \mathbf{x}_0 + V$ for some $\mathbf{x}_0 \in H$ (Exercise 1.153). There are two cases to consider.

Case 1: $\mathbf{x}_0 \notin V$. For every $\mathbf{x} \in X$, there exists unique $\alpha_{\mathbf{x}} \in \Re$ such

$$\mathbf{x} = \alpha_{\mathbf{x}}\mathbf{x}_0 + v \text{ for some } v \in V$$

Define $f(\mathbf{x}) = \alpha_{\mathbf{x}}$. Then $f\colon X \to \Re$. It is straightforward to show that f is linear.

Since $H = \mathbf{x}_0 + V$, $\alpha_{\mathbf{x}} = 1$ if and only if $\mathbf{x} \in H$. Therefore

$$H = \{\, \mathbf{x} \in X : f(\mathbf{x}) = 1 \,\}$$

Case 2: $\mathbf{x}_0 \in V$. In this case, choose some $\mathbf{x}_1 \notin V$. Again, for every $\mathbf{x} \in X$, there exists a unique $\alpha_{\mathbf{x}} \in \Re$ such

$$\mathbf{x} = \alpha_{\mathbf{x}}\mathbf{x}_1 + v \text{ for some } v \in V$$

and $f(\mathbf{x}) = \alpha_{\mathbf{x}}$ is a linear functional on X. Furthermore $\mathbf{x}_0 \in V$ implies $H = V$ (Exercise 1.153) and therefore $f(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in H$. Therefore

$$H = \{\, \mathbf{x} \in X : f(\mathbf{x}) = 0 \,\}$$

Conversely, let f be a nonzero linear functional in X'. Let $V = \text{kernel } f$ and choose $\mathbf{x}_0 \in f^{-1}(1)$. (This is why we require $f \neq \mathbf{0}$). For any $\mathbf{x} \in X$

$$f\left(\mathbf{x} - f(\mathbf{x})\mathbf{x}_0\right) = f(\mathbf{x}) - f(\mathbf{x}) \times 1 = 0$$

so that $\mathbf{x} - f(\mathbf{x})\mathbf{x}_0 \in V$. That is, $\mathbf{x} = f(\mathbf{x})\mathbf{x}_0 + v$ for some $v \in V$. Therefore, $X = \text{lin}\,(\mathbf{x}_0, V)$ so that V is a maximal proper subspace.

For any $c \in \Re$, let $\mathbf{x}_1 \in f^{-1}(c)$. Then, for every $\mathbf{x} \in f^{-1}(c)$, $f(\mathbf{x} - \mathbf{x}_1) = 0$ and

$$\{\, \mathbf{x} : f(\mathbf{x}) = c\} = \{\mathbf{x} : f(\mathbf{x} - \mathbf{x}_1) = 0\,\} = \mathbf{x}_1 + V$$

which is a hyperplane.

3.50 By the previous exercise, there exists a linear functional g such that

$$H = \{\, x \in X : f(x) = c \,\}$$

for some $c \in \Re$. Since $\mathbf{0} \notin H$, $c \neq 0$. Without loss of generality, we can assume that $c = 1$. (Otherwise, take the linear functional $\frac{1}{c}f$).

To show that f is unique, assume that g is another linear functional with

$$H = \{\, \mathbf{x} : f(x) = 1\} = \{\mathbf{x} : g(x) = 1\,\}$$

Then

$$H \subseteq \{\, \mathbf{x} : f(x) - g(x) = 0 \,\}$$

Since H is a maximal subset, X is the smallest subspace containing H. Therefore $f(x) = g(x)$ for every $x \in X$.

3.51 By Exercise 3.49, there exists a linear functional f such that

$$H = \{ x \in X : f(x) = 0 \}$$

Since $\mathbf{x}_0 \notin H$, $f(\mathbf{x}_0) \neq 0$. Without loss of generality, we can normalize so that $f(\mathbf{x}_0) = 1$. (If $f(\mathbf{x}_0) = c \neq 1$, then the linear functional $f' = {}^1\!/_c f$ has $f'(\mathbf{x}_0) = 1$ and kernel $f' = H$.)

To show that f is unique, suppose that g is another linear functional with kernel $g = H$ and $g(\mathbf{x}_0) = 1$. For any $\mathbf{x} \in X$, there exists $\alpha \in \Re$ such that

$$\mathbf{x} = \alpha \mathbf{x}_0 + \mathbf{v}$$

with $v \in H$ (Exercise 1.153). Since $f(\mathbf{v}) = g(\mathbf{v}) = 0$ and $f(\mathbf{x}_0) = g(\mathbf{x}_0) = 1$

$$g(\mathbf{x}) = g(\alpha\mathbf{x}_0 + \mathbf{v}) = \alpha g(x_0) = \alpha f(\mathbf{x}_0) = f(\alpha \mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x})$$

3.52 Assume $f = \lambda g$, $\lambda \neq 0$. Then

$$f(x) = 0 \iff g(x) = 0$$

Conversely, let $H = f^{-1}(0) = g^{-1}(0)$. If $H = X$, then $f = g = \mathbf{0}$. Otherwise, H is a hyperplane containing $\mathbf{0}$. Choose some $\mathbf{x}_0 \notin H$. Every $\mathbf{x} \in X$ has a unique representation $\mathbf{x} = \alpha\mathbf{x}_0 + \mathbf{v}$ with $\mathbf{v} \in H$ (Exercise 1.153) and

$$f(\mathbf{x}) = \alpha f(\mathbf{x}_0)$$
$$g(\mathbf{x}) = \alpha g(\mathbf{x}_0)$$

Let $\lambda = f(\mathbf{x}_0)/g(\mathbf{x}_0)$ so that $f(\mathbf{x}_0) = \lambda g(\mathbf{x}_0)$. Substituting

$$f(\mathbf{x}) = \alpha f(\mathbf{x}_0) = \alpha\lambda g(\mathbf{x}_0) = \lambda g(\mathbf{x})$$

3.53 f continuous implies that the set $\{ x \in X : f(x) = c \} = f^{-1}(c)$ is closed for every $c \in \Re$ (Exercise 2.67). Conversely, let $c = 0$ and assume that $H = \{ x \in X : f(x) = 0 \}$ is closed. There exists $\mathbf{x}_0 \neq \mathbf{0}$ such that $X = \text{lin}\, \{x_0, H\}$ (Exercise 1.153). Let $\mathbf{x}^n \to \mathbf{x}$ be a convergent sequence in X. Then there exist α^n, $\alpha \in \Re$ and $\mathbf{v}^n, v \in H$ such that $\mathbf{x}^n = \alpha^n \mathbf{x}_0 + \mathbf{v}^n$, $\mathbf{x} = \alpha\mathbf{x}_0 + \mathbf{v}$ and

$$\begin{aligned}\|\mathbf{x}^n - \mathbf{x}\| &= \|\alpha^n\mathbf{x}_0 + \mathbf{v}^n - \alpha\mathbf{x}_0 + \mathbf{v}\| \\ &= \|\alpha^n\mathbf{x}_0 - \alpha\mathbf{x}_0 + \mathbf{v}^n - \mathbf{v}\| \\ &\leq |\alpha^n - \alpha|\,\|\mathbf{x}_0\| + \|\mathbf{v}^n - \mathbf{v}\| \\ &\to 0\end{aligned}$$

which implies that $\alpha^n \to \alpha$. By linearity

$$f(\mathbf{x}^n) = \alpha^n f(\mathbf{x}_0) + f(\mathbf{v}^n) = \alpha^n f(\mathbf{x}_0)$$

since $\mathbf{v}^n \in H$ and therefore

$$f(\mathbf{x}^n) = \alpha^n f(\mathbf{x}_0) \to \alpha f(\mathbf{x}_0) = f(\mathbf{x})$$

f is continuous.

3.54

$$\begin{aligned} f(\mathbf{x}+\mathbf{x}',\mathbf{y}) &= \sum_{i=1}^m\sum_{j=1}^n a_{ij}(x_i+x_i')y_j \\ &= \sum_{i=1}^m\sum_{j=1}^n a_{ij}x_iy_j + \sum_{i=1}^m\sum_{j=1}^n a_{ij}x_i'y_j \\ &= f(\mathbf{x},\mathbf{y}) + f(\mathbf{x}',\mathbf{y}) \end{aligned}$$

Similarly, we can show that

$$f(\mathbf{x},\mathbf{y}+\mathbf{y}') = f(\mathbf{x},\mathbf{y}) + f(\mathbf{x},\mathbf{y}')$$

and

$$f(\alpha\mathbf{x},\mathbf{y}) = \alpha f(\mathbf{x},\mathbf{y}) = f(\mathbf{x},\alpha\mathbf{y}) \text{ for every } \alpha \in \Re$$

3.55 Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ be a basis for X and $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ be a basis for Y. Let the numbers a_{ij} represent the action of f on these bases, that is

$$a_{ij} = f(\mathbf{x}_i, \mathbf{y}_j) \qquad i = 1, 2, \dots, m, j = 1, 2, \dots, n$$

and let A be the $m \times n$ matrix of numbers a_{ij}.

Choose any $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ and let their representations in terms of the bases be

$$\mathbf{x} = \sum_{i=1}^m \alpha_i\mathbf{x}_i \text{ and } \mathbf{y} = \sum_{j=1}^n \beta_i\mathbf{y}_j$$

respectively. By the bilinearity of f

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= f(\sum_i \alpha_i\mathbf{x}_i, \sum_j \alpha_j\mathbf{y}_j) \\ &= \sum_i \alpha_i f(\mathbf{x}_i, \sum_j \alpha_j\mathbf{y}_j) \\ &= \sum_i \alpha_i \sum_j \alpha_j f(\mathbf{x}_i,\mathbf{y}_j) \\ &= \sum_i \alpha_i \sum_j \alpha_j a_{ij} \\ &= \sum_i \alpha_i A\mathbf{y} \\ &= \mathbf{x}'A\mathbf{y} \end{aligned}$$

3.56 Every $\mathbf{y} \in X'$ is a linear functional on X. Hence

$$\begin{aligned} \mathbf{y}(\mathbf{x}+\mathbf{x}') &= \mathbf{y}(\mathbf{x}) + \mathbf{y}(\mathbf{x}') \\ \mathbf{y}(\alpha\mathbf{x}) &= \alpha\mathbf{y}(\mathbf{x}) \end{aligned}$$

and therefore

$$\begin{aligned} f(\mathbf{x}+\mathbf{x}',\mathbf{y}) &= \mathbf{y}(\mathbf{x}+\mathbf{x}') = \mathbf{y}(\mathbf{x}) + \mathbf{y}(\mathbf{x}') = f(\mathbf{x},\mathbf{y}) + f(\mathbf{x}',\mathbf{y}) \\ f(\alpha\mathbf{x},\mathbf{y}) &= \mathbf{y}(\alpha\mathbf{x}) = \alpha\mathbf{y}(\mathbf{x}) = \alpha f(\mathbf{x},\mathbf{y}) \end{aligned}$$

In the dual space X'

$$\begin{aligned}(\mathbf{y}+\mathbf{y}')(\mathbf{x}) &\equiv \mathbf{y}(\mathbf{x})+\mathbf{y}'(\mathbf{x})\\ (\alpha\mathbf{y})(\mathbf{x}) &\equiv \alpha\mathbf{y}(\mathbf{x})\end{aligned}$$

and therefore

$$\begin{aligned}f(\mathbf{x},\mathbf{y}+\mathbf{y}') &= (\mathbf{y}+\mathbf{y}')(\mathbf{x}) = \mathbf{y}(\mathbf{x})+\mathbf{y}'(\mathbf{x}) = f(\mathbf{x},\mathbf{y})+f(\mathbf{x},\mathbf{y}')\\ f(\mathbf{x},\alpha\mathbf{y}) &= (\alpha\mathbf{y})(\mathbf{x}) = \alpha\mathbf{y}(\mathbf{x}) = \alpha f(\mathbf{x},\mathbf{y})\end{aligned}$$

3.57 Assume $f_1, f_2 \in BiL(X \times Y, Z)$. Define the mapping $f_1 + f_2 \colon X \times Y \to Z$ by

$$(f_1+f_2)(\mathbf{x},\mathbf{y}) = f_1(\mathbf{x},\mathbf{y}) + f_2(\mathbf{x},\mathbf{y})$$

We have to confirm that $f_1 + f_2$ is bilinear, that is

$$\begin{aligned}(f_1+f_2)(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y}) &= f_1(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y}) + f_2(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y})\\ &= f_1(\mathbf{x}_1,\mathbf{y}) + f_1(\mathbf{x}_2,\mathbf{y}) + f_2(\mathbf{x}_1,\mathbf{y}) + f_2(\mathbf{x}_2,\mathbf{y})\\ &= f_1(\mathbf{x}_1,\mathbf{y}) + f_2(\mathbf{x}_1,\mathbf{y}) + f_1(\mathbf{x}_1,\mathbf{y}) + f_2(\mathbf{x}_2,\mathbf{y})\\ &= (f_1+f_2)(\mathbf{x}_1,\mathbf{y}) + (f_1+f_2)(\mathbf{x}_2,\mathbf{y})\end{aligned}$$

Similarly, we can show that

$$(f_1+f_2)(\mathbf{x},\mathbf{y}_1+\mathbf{y}_2) = (f_1+f_2)(\mathbf{x},\mathbf{y}_1) + (f_1+f_2)(\mathbf{x},\mathbf{y}_2)$$

and

$$(f_1+f_2)(\alpha\mathbf{x},\mathbf{y}) = \alpha(f_1+f_2)(\mathbf{x},\mathbf{y}) = (f_1+f_2)(\mathbf{x},\alpha\mathbf{y})$$

For every $f \in BiL(X \times Y, Z)$ define the function $\alpha f \colon X \times Y \to Z$ by

$$(\alpha f)(\mathbf{x},\mathbf{y}) = \alpha f(\mathbf{x},\mathbf{y})$$

αf is also bilinear, since

$$\begin{aligned}(\alpha f)(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y}) &= \alpha f(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y})\\ &= \alpha f(\mathbf{x}_1,\mathbf{y}) + \alpha f(\mathbf{x}_2,\mathbf{y})\\ &= (\alpha f)(\mathbf{x}_1,\mathbf{y}) + (\alpha f(\mathbf{x}_2,\mathbf{y})\end{aligned}$$

Similarly

$$\begin{aligned}(\alpha f)(\mathbf{x},\mathbf{y}_1+\mathbf{y}_2) &= (\alpha f)(\mathbf{x},\mathbf{y}_1) + (\alpha f)(\mathbf{x},\mathbf{y}_2)\\ (\alpha f)(\beta\mathbf{x},\mathbf{y}) &= \beta(\alpha f)(\mathbf{x},\mathbf{y}) = (\alpha f)(\mathbf{x},\beta\mathbf{y})\end{aligned}$$

Analogous to (Exercise 2.78), $f_1 + f_2$ and αf are also continuous

3.58 1. $BL(Y, Z)$ is a linear space and therefore so is $BL(X, BL(Y, Z))$ (Exercise 3.33).

2. $\varphi_{\mathbf{x}}$ is linear and therefore

$$f(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = \varphi(\mathbf{x})(\mathbf{y}_1 + \mathbf{y}_2) = \varphi(\mathbf{x})(\mathbf{y}_1) + \varphi(\mathbf{x})(\mathbf{y}_2) = f(\mathbf{x}, \mathbf{y}_1) + f(\mathbf{x}, \mathbf{y}_2)$$

and

$$f(\mathbf{x}, \alpha\mathbf{y}) = \varphi(\mathbf{x})(\alpha\mathbf{y}) = \alpha\varphi(\mathbf{x})(\mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$$

Similarly, φ is linear and therefore

$$f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = \varphi_{\mathbf{x}_1+\mathbf{x}_2}(\mathbf{y}) = \varphi_{\mathbf{x}_1}(\mathbf{y}) + \varphi_{\mathbf{x}_2}(\mathbf{y}) = f(\mathbf{x}_1, \mathbf{y}) + f(\mathbf{x}_2, \mathbf{y})$$

and

$$f(\alpha\mathbf{x}, \mathbf{y}) = \varphi_{\alpha\mathbf{x}}(\mathbf{y}) = \alpha\varphi_{\mathbf{x}}(\mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$$

f is bilinear

3. Let $f \in BiL(X \times Y, Z)$. For every $\mathbf{x} \in X$, the partial function $f_{\mathbf{x}} \colon Y \to Z$ is linear. Therefore $f_{\mathbf{x}} \in BL(Y, Z)$ and $\varphi \in BL(X, BL(Y, Z))$.

3.59 Bilinearity and symmetry imply

$$\begin{aligned} f(\mathbf{x} - \alpha\mathbf{y}, \mathbf{x} - \alpha\mathbf{y}) &= f(\mathbf{x}, \mathbf{x} - \alpha\mathbf{y}) - \alpha f(\mathbf{y}, \mathbf{x} - \alpha\mathbf{y}) \\ &= f(\mathbf{x}, \mathbf{x}) - \alpha f(\mathbf{x}, \mathbf{y}) - \alpha f(\mathbf{y}, \mathbf{x}) + \alpha^2 f(\mathbf{y}, \mathbf{y}) \\ &= f(\mathbf{x}, \mathbf{x}) - 2\alpha f(\mathbf{x}, \mathbf{y}) + \alpha^2 f(\mathbf{y}, \mathbf{y}) \end{aligned}$$

Nonnegativity implies

$$f(\mathbf{x} - \alpha\mathbf{y}, \mathbf{x} - \alpha\mathbf{y}) = f(\mathbf{x}, \mathbf{x}) - 2\alpha f(\mathbf{x}, \mathbf{y}) + \alpha^2 f(\mathbf{y}, \mathbf{y}) \geq 0 \tag{3.1}$$

for every $\mathbf{x}, \mathbf{y} \in X$ and $\alpha \in \Re$

Case 1 $f(\mathbf{x}, \mathbf{x}) = f(\mathbf{y}, \mathbf{y}) = 0$ Then (3.1) becomes

$$-2\alpha f(\mathbf{x}, \mathbf{y}) \geq 0$$

Setting $\alpha = f(\mathbf{x}, \mathbf{y})$ generates

$$-2\big(f(\mathbf{x}, \mathbf{y})\big)^2 \geq 0$$

which implies that

$$f(\mathbf{x}, \mathbf{y}) = 0$$

Case 2 Either $f(\mathbf{x}, \mathbf{x}) > 0$ or $f(\mathbf{y}, \mathbf{y}) > 0$. Without loss of generality, assume $f(\mathbf{y}, \mathbf{y}) > 0$ and set $\alpha = f(\mathbf{x}, \mathbf{y})/f(\mathbf{y}, \mathbf{y})$ in (3.1). That is

$$f(\mathbf{x}, \mathbf{x}) - 2\left(\frac{f(\mathbf{x}, \mathbf{y})}{f(\mathbf{y}, \mathbf{y})}\right) f(\mathbf{x}, \mathbf{y}) + \left(\frac{f(\mathbf{x}, \mathbf{y})}{f(\mathbf{y}, \mathbf{y})}\right)^2 f(\mathbf{y}, \mathbf{y}) \geq 0$$

or

$$f(\mathbf{x}, \mathbf{x}) - \frac{f(\mathbf{x}, \mathbf{y})^2}{f(\mathbf{y}, \mathbf{y})} \geq 0$$

which implies

$$\big(f(\mathbf{x}, \mathbf{y})\big)^2 \leq f(\mathbf{x}, \mathbf{x}) f(\mathbf{y}, \mathbf{y}) \text{ for every } \mathbf{x}, \mathbf{y} \in X$$

3.60 A Euclidean space is a finite-dimensional normed space, which is complete (Proposition 1.4).

3.61 $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y}$ satisfies the requirements of Exercise 3.59 and therefore

$$(\mathbf{x}^T\mathbf{y})^2 \leq (\mathbf{x}^T\mathbf{x})(\mathbf{y}^T\mathbf{y})$$

Taking square roots

$$\left|\mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{x}\| \; \|\mathbf{y}\|$$

3.62 By definition, the inner product is a bilinear functional. To show that it is continuous, let X be an inner product space with inner product denote by $\mathbf{x}^T\mathbf{y}$. Let $\mathbf{x}^n \to \mathbf{x}$ and $\mathbf{y}^n \to \mathbf{y}$ be sequences in X.

$$\begin{aligned}\left|(\mathbf{x}^n)^T\mathbf{y}^n - \mathbf{x}^T\mathbf{y}\right| &= \left|(\mathbf{x}^n)^T\mathbf{y}^n - (\mathbf{x}^n)^T\mathbf{y} + (\mathbf{x}^n)^T\mathbf{y} - \mathbf{x}^T\mathbf{y}\right| \\ &\leq \left|(\mathbf{x}^n)^T\mathbf{y}^n - (\mathbf{x}^n)^T\mathbf{y}\right| + \left|(\mathbf{x}^n)^T\mathbf{y} - \mathbf{x}^T\mathbf{y}\right| \\ &\leq \left|(\mathbf{x}^n)^T(\mathbf{y}^n - \mathbf{y})\right| + \left|(\mathbf{x}^n - \mathbf{x})^T\mathbf{y}\right|\end{aligned}$$

Applying the Cauchy-Schwartz inequality

$$\left|(\mathbf{x}^n)^T\mathbf{y}^n - \mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{x}^n\| \; \|\mathbf{y}^n - \mathbf{y}\| + \|\mathbf{x}^n - \mathbf{x}\| \; \|\mathbf{y}\|$$

Since the sequence $\mathbf{x}^n$ converges, it is bounded, that is there exists M such that $\|\mathbf{x}^n\| \leq M$ for every n. Therefore

$$\left|(\mathbf{x}^n)^T\mathbf{y}^n - \mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{x}^n\| \, \|\mathbf{y}^n - \mathbf{y}\| + \|\mathbf{x}^n - \mathbf{x}\| \, \|\mathbf{y}\| \leq M \, \|\mathbf{y}^n - \mathbf{y}\| + \|\mathbf{x}^n - \mathbf{x}\| \, \|\mathbf{y}\| \to \mathbf{0}$$

3.63 Applying the properties of the inner product

- $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} \geq 0$
- $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} = 0$ if and only if $\mathbf{x} = 0$
- $\|\alpha\mathbf{x}\| = \sqrt{(\alpha\mathbf{x})^T(\alpha\mathbf{x})} = \sqrt{\alpha^2\mathbf{x}^T\mathbf{x}} = |\alpha| \, \|\mathbf{x}\|$

To prove the triangle inequality, observe that bilinearity and symmetry imply

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y})^T(\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x}^T\mathbf{x} + \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{x} + \mathbf{z}^T\mathbf{z} \\ &= \mathbf{x}^T\mathbf{x} + 2\mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{y} \\ &= \|\mathbf{x}\|^2 + 2\mathbf{x}^T\mathbf{y} + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\left|\mathbf{x}^T\mathbf{y}\right| + \|\mathbf{y}\|^2\end{aligned}$$

Applying the Cauchy-Schwartz inequality

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|^2 &\leq \|\mathbf{x}\|^2 + 2\,\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2\end{aligned}$$

3.64 For every $\mathbf{y} \in X$, the partial function $f_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}^T\mathbf{y}$ is a linear functional on X (since $\mathbf{x}^T\mathbf{y}$ is bilinear). Continuity follows from the Cauchy-Schwartz inequality, since for every $\mathbf{x} \in X$

$$|f_{\mathbf{y}}(\mathbf{x})| = \left|\mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{y}\| \, \|\mathbf{x}\|$$

which shows that $\|f_{\mathbf{y}}\| \leq \|\mathbf{y}\|$. In fact, $\|f_{\mathbf{y}}\| = \|\mathbf{y}\|$ since

$$\begin{aligned}
\|f_{\mathbf{y}}\| &= \sup_{\|\mathbf{x}\|=1} |f_{\mathbf{y}}(\mathbf{x})| \\
&\geq \left| f_{\mathbf{y}}\left(\frac{\mathbf{y}}{\|\mathbf{y}\|}\right) \right| \\
&= \left| \left(\frac{\mathbf{y}}{\|\mathbf{y}\|}\right)^T \mathbf{y} \right| \\
&= \frac{1}{\|\mathbf{y}\|}\mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|
\end{aligned}$$

3.65 By the Weierstrass Theorem (Theorem 2.2), the continuous function $g(\mathbf{x}) = \|\mathbf{x}\|$ attains a maximum on the compact set S at some point $\mathbf{x}_0$.

We claim that $\mathbf{x}_0$ is an extreme point. Suppose not. Then, there exist $\mathbf{x}_1, \mathbf{x}_2 \in S$ such that

$$\mathbf{x}_0 = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 = \mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)$$

Since $\mathbf{x}_0$ maximizes $\|\mathbf{x}\|$ on S

$$\begin{aligned}
\|\mathbf{x}_2\|^2 \leq \|\mathbf{x}_0\|^2 &= \big(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)\big)^T \big(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)\big) \\
&= \|\mathbf{x}_2\|^2 + 2\alpha\mathbf{x}_2^T(\mathbf{x}_1 - \mathbf{x}_2) + \alpha^2 \|\mathbf{x}_1 - \mathbf{x}_2\|^2
\end{aligned}$$

or

$$2\mathbf{x}_2^T(\mathbf{x}_1 - \mathbf{x}_2) + \alpha \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \geq 0 \tag{3.2}$$

Similarly, interchanging the role of $\mathbf{x}_1$ and $\mathbf{x}_2$

$$2\mathbf{x}_1^T(\mathbf{x}_2 - \mathbf{x}_1) + \alpha \|\mathbf{x}_2 - \mathbf{x}_1\|^2 \geq 0$$

or

$$-2\mathbf{x}_1^T(\mathbf{x}_1 - \mathbf{x}_2) + \alpha \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \geq 0 \tag{3.3}$$

Adding the inequalities (3.2) and (3.3) yields

$$2(\mathbf{x}_2 - \mathbf{x}_1)^T(\mathbf{x}_1 - \mathbf{x}_2) + 2\alpha \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \geq 0$$

or

$$2(\mathbf{x}_2 - \mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) = -2(\mathbf{x}_2 - \mathbf{x}_1)^T(\mathbf{x}_1 - \mathbf{x}_2) \leq 2\alpha \|\mathbf{x}_1 - \mathbf{x}_2\|^2$$

and therefore

$$\|\mathbf{x}_2 - \mathbf{x}_1\| \leq \alpha \|\mathbf{x}_2 - \mathbf{x}_1\|$$

Since $0 < \alpha < 1$, this implies that $\|\mathbf{x}_1 - \mathbf{x}_2\| = 0$ or $\mathbf{x}_1 = \mathbf{x}_2$ which contradicts our premise that $\mathbf{x}_0$ is not an extreme point.

3.66 Using bilinearity and symmetry of the inner product

$$\begin{aligned}
\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y})^T(\mathbf{x} + \mathbf{y}) + (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y}) \\
&= \mathbf{x}^T\mathbf{x} + \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{x} + \mathbf{y}^T\mathbf{y} + \\
&\quad \mathbf{x}^T\mathbf{x} - \mathbf{x}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} + \mathbf{y}^T\mathbf{y} \\
&= 2\mathbf{x}^T\mathbf{x} + 2\mathbf{y}^T\mathbf{y} \\
&= 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2
\end{aligned}$$

3.67 Note that $\|x\| = \|y\| = 1$ and

$$\|x + y\| = \sup_{0 \leq t \leq 1} \big(x(t) + y(t)\big) = \sup_{0 \leq t \leq 1} (1 + t) = 2$$
$$\|x - y\| = \sup_{0 \leq t \leq 1} \big(x(t) - y(t)\big) = \sup_{0 \leq t \leq 1} (1 - t) = 1$$

so that

$$\|x + y\|^2 + \|x - y\|^2 = 5 \neq 2\|x\|^2 + 2\|x\|^2$$

Since $\mathbf{x}$ and $\mathbf{y}$ do not satisfy the Parallelogram Law (Exercise 3.66), $C(X)$ cannot be an inner product space.

3.68 Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a set of pairwise orthogonal vectors. Assume

$$\mathbf{0} = \alpha\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n$$

Using bilinearity, this implies

$$0 = \mathbf{0}^T\mathbf{x}_j = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i^T\mathbf{x}_j = \alpha_j \|\mathbf{x}_j\|$$

for every $j = 1, 2, \ldots, n$. Since $\mathbf{x}_j \neq \mathbf{0}$, this implies $\alpha_j = 0$ for every $j = 1, 2, \ldots, n$. We conclude that the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is linearly independent (Exercise 1.133).

3.69 Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be a orthonormal basis for X. Since A represents f

$$f(\mathbf{x}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{x}_i$$

for $j = 1, 2, \ldots, n$. Taking the inner product with $\mathbf{x}_i$,

$$\mathbf{x}_i^T f(\mathbf{x}_j) = \mathbf{x}_i^T \left(\sum_{i=1}^{n} a_{ij}\mathbf{x}_i \right) = \sum_{i=1}^{n} a_{ij}\mathbf{x}_i^T\mathbf{x}_j$$

Since $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is orthonormal

$$\mathbf{x}_k^T\mathbf{x}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

so that the last sum simplifies to

$$\mathbf{x}_i^T f(\mathbf{x}_j) = a_{ij} \text{ for every } \; i, j$$

3.70 1. By the Cauchy-Schwartz inequality

$$\left|\mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{x}\| \, \|\mathbf{x}\|$$

for every $\mathbf{x}$ and $\mathbf{y}$, so that

$$|\cos\theta| = \left| \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|} \right| \leq 1$$

which implies

$$-1 \leq \cos\theta \leq 1$$

2. Since $\cos 90 = 0$, $\theta = 90$ implies that $\mathbf{x}^T\mathbf{y} = 0$ or $\mathbf{x} \perp \mathbf{y}$. Conversely, if $\mathbf{x} \perp \mathbf{y}$, $\mathbf{x}^T\mathbf{y} = 0$ and $\cos\theta = 0$ which implies $\theta = 90$ degrees.

3.71 By bilinearity

$$\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y})^T(\mathbf{x}+\mathbf{y}) = \|\mathbf{x}\|^2 + \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{x} + \|\mathbf{y}\|^2$$

If $\mathbf{x} \perp \mathbf{y}$, $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x} = 0$ and

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

3.72 1. Choose some $\hat{\mathbf{x}} \in S$ and let $\hat{S}$ be the set of all $\mathbf{x} \in S$ which are closer to $\mathbf{y}$ than $\hat{\mathbf{x}}$, that is

$$\hat{S} = \{\, \mathbf{x} \in S : \|\mathbf{x}-\mathbf{y}\| \le \|\hat{\mathbf{x}}-\mathbf{y}\| \,\}$$

$\hat{S}$ is compact (Proposition 1.4).

By the Weierstrass Theorem (Theorem 2.2), the continuous function $g(\mathbf{x}) = \|\mathbf{x}\mathbf{y}\|$ attains a minimum on $\hat{S}$ at some point $\mathbf{x}_0 \in \hat{S}$. That is

$$\|\mathbf{x}_0 - \mathbf{y}\| \le \|\mathbf{x}-\mathbf{y}\| \text{ for every } \mathbf{x} \in \hat{S}$$

A fortiori

$$\|\mathbf{x}_0 - \mathbf{y}\| \le \|\mathbf{x}-\mathbf{y}\| \text{ for every } \mathbf{x} \in S$$

2. Suppose there exists some $\mathbf{x}_1 \in S$ such that

$$\|\mathbf{x}_1 - \mathbf{y}\| = \|\mathbf{x}_0 - \mathbf{y}\| = \delta$$

By the Parallelogram Law (Exercise 3.66)

$$\begin{aligned}
\|\mathbf{x}_0 - \mathbf{x}_1\|^2 &= \|\mathbf{x}_0 - \mathbf{y} + \mathbf{y} - \mathbf{x}_1\|^2 \\
&= 2\|\mathbf{x}_0 - \mathbf{y}\|^2 + 2\|\mathbf{x}_1 - \mathbf{y}\|^2 - \|(\mathbf{x}_0 - \mathbf{y}) - (\mathbf{y} - \mathbf{x}_1)\|^2 \\
&= 2\|\mathbf{x}_0 - \mathbf{y}\|^2 + 2\|\mathbf{x}_1 - \mathbf{y}\|^2 - 2^2\left\|\frac{1}{2}(\mathbf{x}_0 + \mathbf{x}_1) - \mathbf{y}\right\|^2 \\
&= 2\delta^2 + 2\delta^2 - 2^2\left\|\frac{1}{2}(\mathbf{x}_0 + \mathbf{x}_1) - \mathbf{y}\right\|^2
\end{aligned}$$

since $\frac{1}{2}(\mathbf{x}_0 + \mathbf{x}_1) \in S$ and therefore $\left\|\frac{1}{2}(\mathbf{x}_0 + \mathbf{x}_1) - \mathbf{y}\right\| \ge \delta$ so that

$$\|\mathbf{x}_0 - \mathbf{x}_1\|^2 \le 2\delta^2 + 2\delta^2 - 4\delta^2 = 0$$

which implies that $\mathbf{x}_1 = \mathbf{x}_0$.

3. Let $\mathbf{x} \in S$. Since S is convex, the line segment $\alpha\mathbf{x} + (1-\alpha)\mathbf{x}_0 = \mathbf{x}_0 + \alpha(\mathbf{x}-\mathbf{x}_0) \in S$ and therefore (since $\mathbf{x}_0$ is the closest point)

$$\begin{aligned}
\|\mathbf{x}_0 - \mathbf{y}\|^2 &\le \left\|\big(\mathbf{x}_0 + \alpha(\mathbf{x}-\mathbf{x}_0)\big) - \mathbf{y}\right\|^2 \\
&= \|(\mathbf{x}_0 - \mathbf{y}) + \alpha(\mathbf{x}-\mathbf{x}_0)\|^2 \\
&= \big((\mathbf{x}_0 - \mathbf{y}) + \alpha(\mathbf{x}-\mathbf{x}_0)\big)^T\big((\mathbf{x}_0 - \mathbf{y}) + \alpha(\mathbf{x}-\mathbf{x}_0)\big) \\
&= \|\mathbf{x}_0 - \mathbf{y}\|^2 + 2\alpha(\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x}-\mathbf{x}_0) + \alpha^2\|\mathbf{x}-\mathbf{x}_0\|^2
\end{aligned}$$

which implies that

$$2\alpha(\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x} - \mathbf{x}_0) + \alpha^2 \|\mathbf{x} - \mathbf{x}_0\|^2 \geq 0$$

Dividing through by α

$$2(\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x} - \mathbf{x}_0) + \alpha \|\mathbf{x} - \mathbf{x}_0\|^2 \geq 0$$

which inequality must hold for every $0 < \alpha < 1$. Letting $\alpha \to 0$, we must have

$$(\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x} - \mathbf{x}_0) \geq 0$$

as required.

3.73 1. Using the Parallelogram Law (Exercise 3.66),

$$\begin{aligned}\|\mathbf{x}^m - \mathbf{x}^n\|^2 &= \|(\mathbf{x}^m - \mathbf{y}) + (\mathbf{y} - \mathbf{x}^n)\|^2 \\ &= 2\|\mathbf{x}^m - \mathbf{y}\|^2 + 2\|\mathbf{y} - \mathbf{x}^n\|^2 - 2\|\mathbf{x}^m + \mathbf{x}^n\|^2\end{aligned}$$

for every m, n. Since S is convex, $(\mathbf{x}^m + \mathbf{x}^n)/2 \in S$ and therefore $\|\mathbf{x}^m + \mathbf{x}^n\| \geq 2d$. Therefore

$$\|\mathbf{x}^m - \mathbf{x}^n\|^2 = 2\|\mathbf{x}^m - \mathbf{y}\|^2 + 2\|\mathbf{y} - \mathbf{x}^n\|^2 - 4d^2$$

Since $\|\mathbf{x}^m - \mathbf{y}\| \to d$ and $\|\mathbf{x}^n - \mathbf{y}\| \to d$ as $m, n \to \infty$, we conclude that $\|\mathbf{x}^m - \mathbf{x}^n\|^2 \to 0$. That is, $(\mathbf{x}^n)$ is a Cauchy sequence.

2. Since S is a closed subspace of complete space, there exists $\mathbf{x}_0 \in S$ such that $\mathbf{x}^n \to \mathbf{x}_0$. By continuity of the norm

$$\|\mathbf{x}_0 - \mathbf{y}\| = \lim_{n \to \infty} \|\mathbf{x}^n - \mathbf{y}\| = d$$

Therefore

$$\|\mathbf{x}_0 - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| \text{ for every } \mathbf{x} \in S$$

Uniqueness follows in the same manner as the finite-dimensional case.

3.74 Define $g\colon T \to S$ by

$$g(\mathbf{y}) = \{\, \mathbf{x} \in S : \mathbf{x} \text{ is closest to } \mathbf{y} \,\}$$

The function g is well-defined since for every $\mathbf{y} \in T$ there exists a unique point $\mathbf{x} \in S$ which is closest to $\mathbf{y}$ (Exercise 3.72). Clearly, for every $\mathbf{x} \in S$, $\mathbf{x}$ is the closest point to $\mathbf{x}$. Therefore $g(\mathbf{x}) = \mathbf{x}$ for every $\mathbf{x} \in S$.

To show that g is continuous, choose any $\mathbf{y}_1$ and $\mathbf{y}_2$ in T

$$\mathbf{x}_1 = g(\mathbf{y}_1) \text{ and } \mathbf{x}_2 = g(\mathbf{y}_2)$$

be the corresponding closest points in S. Then

$$\begin{aligned}
\|(\mathbf{y}_1 - \mathbf{y}_2) - (\mathbf{x}_1 - \mathbf{x}_2)\|^2 &= \big((\mathbf{y}_1 - \mathbf{y}_2) - (\mathbf{x}_1 - \mathbf{x}_2)\big)^T \big((\mathbf{y}_1 - \mathbf{y}_1) - (\mathbf{x}_1 - \mathbf{x}_2)\big) \\
&= (\mathbf{y}_1 - \mathbf{y}_2)^T(\mathbf{y}_1 - \mathbf{y}_2) + (\mathbf{x}_1 - \mathbf{x}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&\quad - 2(\mathbf{y}_1 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 + \|\mathbf{x}_1 - \mathbf{x}_2\|^2 - 2(\mathbf{y}_1 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 + \|\mathbf{x}_1 - \mathbf{x}_2\|^2 - 2(\mathbf{y}_1 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&\quad - 2\,\|\mathbf{x}_1 - \mathbf{x}_2\|^2 + 2(\mathbf{x}_1 - \mathbf{x}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 - \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \\
&\quad + 2\big((\mathbf{x}_1 - \mathbf{x}_2) - (\mathbf{y}_1 - \mathbf{y}_2)\big)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 - \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \\
&\quad + 2(\mathbf{x}_1 - \mathbf{y}_1)^T(\mathbf{x}_1 - \mathbf{x}_2) - 2(\mathbf{x}_2 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 - \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \\
&\quad - 2(\mathbf{x}_1 - \mathbf{y}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) - 2(\mathbf{x}_2 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2)
\end{aligned}$$

so that

$$\begin{aligned}
\|\mathbf{y}_1 - \mathbf{y}_2\|^2 - \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= \|(\mathbf{y}_1 - \mathbf{y}_2) - (\mathbf{x}_1 - \mathbf{x}_2)\|^2 \\
&\quad + 2(\mathbf{x}_1 - \mathbf{y}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) + 2(\mathbf{x}_2 - \mathbf{y}_2)^T(\mathbf{x}_1 - x_2)
\end{aligned}$$

Using Exercise 3.72

$$(\mathbf{x}_1 - \mathbf{y}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) \geq 0 \text{ and } (\mathbf{x}_2 - \mathbf{y}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \geq 0$$

which implies that the left-hand side

$$\|\mathbf{y}_1 - \mathbf{y}_2\|^2 - \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \geq 0$$

or

$$\|\mathbf{x}_1 - \mathbf{x}_2\| = \|g(\mathbf{y}_1) - g(\mathbf{y}_2)\| \leq \|\mathbf{y}_1 - \mathbf{y}_2\|$$

g is Lipschitz continuous.

3.75 Let $S = \text{kernel } f$. Then S is a closed subspace of X. If $S = X$, then f is the zero functional and $\mathbf{y} = \mathbf{0}$ is the required element. Otherwise chose any $\mathbf{y} \notin S$ and let $\mathbf{x}_0$ be the closest point in S (Exercise 3.72). Define $\mathbf{z} = \mathbf{x}_0 - \mathbf{y}$. Then $\mathbf{z} \neq \mathbf{0}$ and

$$\mathbf{z}^T\mathbf{x} \geq 0 \text{ for every } \mathbf{x} \in S$$

Since S is subspace, this implies that

$$\mathbf{z}^T\mathbf{x} = 0 \text{ for every } \mathbf{x} \in S$$

that is $\mathbf{z}$ is orthogonal to S.

Let $\hat{S}$ be the subset of X defined by

$$\hat{S} = \{\, f(\mathbf{x})\mathbf{z} - f(\mathbf{z})\mathbf{x} : \mathbf{x} \in X \,\}$$

For every $\mathbf{x} \in \hat{S}$

$$f(\mathbf{x}) = f\big(f(\mathbf{x})\mathbf{z} - f(\mathbf{z})\mathbf{x}\big) = f(\mathbf{x})f(\mathbf{z}) - f(\mathbf{z})f(\mathbf{x}) = 0$$

Therefore $\hat{S} \subseteq S$. For every $\mathbf{x} \in X$

$$\big(f(\mathbf{x})\mathbf{z} - f(\mathbf{z})\mathbf{x}\big)^T \mathbf{z} = f(\mathbf{x})\mathbf{z}^T\mathbf{z} - f(\mathbf{z})\mathbf{x}^T\mathbf{z} = 0$$

since $\mathbf{z} \in S^{\perp}$. Therefore

$$f(\mathbf{x}) = \frac{f(\mathbf{z})}{\|\mathbf{z}\|^2}\mathbf{x}^T\mathbf{z} = \mathbf{x}^T\left(\frac{\mathbf{z}f(\mathbf{z})}{\|\mathbf{z}\|^2}\right) = \mathbf{x}^T\mathbf{y}$$

where

$$\mathbf{y} = \frac{\mathbf{z}f(\mathbf{z})}{\|\mathbf{z}\|^2}$$

3.76 X^* is always complete (Proposition 3.3). To show that it is a Hilbert space, we have to that it has an inner product. For this purpose, it will be clearer if we use an alternative notation $<\mathbf{x}, \mathbf{y}>$ to denote the inner product of $\mathbf{x}$ and $\mathbf{y}$. Assume X is a Hilbert space. By the Riesz representation theorem (Exercise 3.75), for every $f \in X^*$ there exists $\mathbf{y}_f \in X$ such that

$$f(\mathbf{x}) = <\mathbf{x}, \mathbf{y}_f> \text{ for every } \mathbf{x} \in X$$

Furthermore, if $\mathbf{y}_f$ represents f and $\mathbf{y}_g$ represents $g \in X^*$, then $\mathbf{y}_f + \mathbf{y}_g$ represents $f + g$ and $\alpha\mathbf{y}_f$ represents αf since

$$\begin{aligned}(f+g)(\mathbf{x}) &= f(\mathbf{x}) + g(\mathbf{x}) = <\mathbf{x}, \mathbf{y}_f> + <\mathbf{x}, \mathbf{y}_g> = <\mathbf{x}, \mathbf{y}_f + \mathbf{y}_g> \\ (\alpha f)(\mathbf{x}) &= \alpha f(\mathbf{x}) = \alpha <\mathbf{x}, \mathbf{y}_f> = <\mathbf{x}, \alpha\mathbf{y}_f>\end{aligned}$$

Define an inner product on X^* by

$$<f, g> = <\mathbf{y}_g, \mathbf{y}_f>$$

We show that it satisfies the properties of an inner product, namely

symmetry $<f, g> = <\mathbf{y}_g, \mathbf{y}_f> = <\mathbf{y}_f, \mathbf{y}_g> = <g, f>$

additivity $<f_1 + f_2, g> = <\mathbf{y}_g, \mathbf{y}_{f_1+f_2}> = <\mathbf{y}_g, \mathbf{y}_{f_1} + \mathbf{y}_{f_2}> = <f_1, g> + <f_2, g>$

homogeneity $<\alpha f, g> = <\mathbf{y}_g, \alpha\mathbf{y}_f> = \alpha <\mathbf{y}_g, \mathbf{y}_f> = \alpha <f, g>$

positive definiteness $<f, g> = <\mathbf{y}_g, \mathbf{y}_f> \geq 0$ and $<f, g> = <\mathbf{y}_g, \mathbf{y}_f> = 0$ if and only if $f = g$.

Therefore, X^* is a complete inner product space, that is a Hilbert space.

3.77 Let X be a Hilbert space. Applying the previous exercise a second time, X^{**} is also a Hilbert space. Let F be an arbitrary functional in X^{**}. By the Riesz representation theorem, there exists $g \in X^*$ such that

$$F(f) = <f, g> \text{ for every } f \in X^*$$

Again by the Riesz representation theorem, there exists $\mathbf{x}_f$ (representing f) and $\mathbf{x}_F$ (representing g) in X such that

$$F(f) = <f, g> = <\mathbf{x}_F, \mathbf{x}_f>$$

and

$$f(\mathbf{x}) = <\mathbf{x}, \mathbf{x}_f>$$

In particular,

$$f(\mathbf{x}_F) =< \mathbf{x}_F, \mathbf{x}_f >= F(f)$$

That is, for every $F \in X^{**}$, there exists an element $\mathbf{x}_F \in X$ such that

$$F(f) = f(\mathbf{x}_F)$$

X is reflexive.

3.78 1. Adapt Exercise 3.64.

2. By Exercise 3.75, there exists unique $\mathbf{x}^* \in X$ such that

$$f_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}^T \mathbf{x}^*$$

3. Substituting

$$f(\mathbf{x})^T \mathbf{y} = f_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}^T \mathbf{x}^* = x^T f^*(\mathbf{y})$$

4. For every $\mathbf{y}_1, \mathbf{y}_2 \in Y$

$$\mathbf{x}^T \left(f^*(\mathbf{y}_1 + \mathbf{y}_2)\right) = f(\mathbf{x})^T \left(\mathbf{y}_1 + \mathbf{y}_2\right) = f(\mathbf{x})^T \mathbf{y}_1 + f(\mathbf{x})^T \mathbf{y}_1 = \mathbf{x}^T f^*(\mathbf{y}_1) + \mathbf{x}^T f^*(\mathbf{y}_1)$$

and for every $\mathbf{y} \in Y$

$$\mathbf{x}^T f^*(\alpha \mathbf{y}) = f(\mathbf{x})^T \alpha \mathbf{y} = \alpha f(\mathbf{x})^T \mathbf{y} = \alpha \mathbf{x}^T f^*(\mathbf{y}) = \mathbf{x}^T \alpha f^*(\mathbf{y})$$

3.79 The zero element $\mathbf{0}_X$ is a fixed point of every linear operator (Exercise 3.13).

3.80

$$AA^{-1} = I$$

so that

$$\det(A)\det(A^{-1}) = \det(I) = 1$$

3.81 Expanding along the ith row using (3.8)

$$\begin{aligned}
\det(C) &= \sum_{j=1}^{n} (-1)^{i+j} (\alpha a_{ij} + \beta b_{ij}) \det(C_{ij}) \\
&= \alpha \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(C_{ij}) + \beta \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(C_{ij})
\end{aligned}$$

But the matrices differ only in the ith row and therefore

$$A_{ij} = B_{ij} = C_{ij}, \qquad j = 1, 2, \ldots n$$

so that

$$\begin{aligned}
\det(C) &= \alpha \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) + \beta \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(B_{ij}) \\
&= \alpha \det(A) + \beta \det(B)
\end{aligned}$$

3.82 Suppose that $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors corresponding to the eigenvalue λ. By linearity

$$f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2) = \lambda \mathbf{x}_1 + \lambda \mathbf{x}_2 = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$$

and

$$f(\alpha \mathbf{x}_1) = \alpha f(\mathbf{x}_1) = \alpha \lambda \mathbf{x}$$

Therefore $\mathbf{x}_1 + \mathbf{x}_2$ and $\alpha \mathbf{x}_1$ are also eigenvectors.

3.83 Suppose f is singular. Then there exists $\mathbf{x} \neq \mathbf{0}$ such that $f(\mathbf{x}) = 0$. Therefore $\mathbf{x}$ is an eigenvector with eigenvalue 0. Conversely, if 0 is an eigenvalue

$$f(\mathbf{x}) = 0\mathbf{x} = 0$$

for any $\mathbf{x} \neq \mathbf{0}$. Therefore f is singular.

3.84 Since $f(\mathbf{x}) = \lambda \mathbf{x}$

$$f(\mathbf{x})^T \mathbf{x} = \lambda \mathbf{x}^T \mathbf{x} = \lambda \mathbf{x}^T \mathbf{x}$$

3.85 By Exercise 3.69

$$\begin{aligned} a_{ij} &= \mathbf{x}_i^T f(\mathbf{x}_j) \\ a_{ji} &= \mathbf{x}_j^T f(\mathbf{x}_i) = f(\mathbf{x}_i)^T \mathbf{x}_j \end{aligned}$$

and therefore

$$a_{ij} = a_{ji} \iff \mathbf{x}_i^T f(\mathbf{x}_j) = f(\mathbf{x}_i)^T \mathbf{x}_j$$

3.86 By bilinearity

$$\mathbf{x}_1^T f(\mathbf{x}_2) = \mathbf{x}_1^T \lambda_2 \mathbf{x}_2 = \lambda_2 \mathbf{x}_1^T \mathbf{x}_2$$

$$f(\mathbf{x}_1)^T \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^T \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^T \mathbf{x}_2$$

Since f is symmetric, this implies

$$(\lambda_1 - \lambda_2)\mathbf{x}_1^T \mathbf{x}_2 = 0$$

and $\lambda_1 \neq \lambda_2$ implies $\mathbf{x}_1^T \mathbf{x}_2 = 0$.

3.87

1. Since S compact and f is continuous (Exercises 3.31, 3.62), the maximum is attained at some $\mathbf{x}_0 \in S$ (Theorem 2.2), that is

   $$\lambda = f(\mathbf{x}_0)^T \mathbf{x}_0 \geq f(\mathbf{x})^T \mathbf{x} \text{ for every } \mathbf{x} \in S$$

   Hence

   $$g(\mathbf{x}, \mathbf{y}) = \big(\lambda \mathbf{x} - f(\mathbf{x})\big)^T \mathbf{y}$$

   is well-defined.

2. For any $\mathbf{x} \in X$

   $$\begin{aligned} g(\mathbf{x}, \mathbf{x}) &= \big(\lambda \mathbf{x} - f(\mathbf{x})\big)^T \mathbf{x} \\ &= \lambda \mathbf{x}^T \mathbf{x} - f(\mathbf{x})^T \mathbf{x} \\ &= \lambda \|\mathbf{x}\|^2 - f(\mathbf{x})^T \mathbf{x} \\ &= \lambda \|\mathbf{x}\|^2 - \|\mathbf{x}\|^2 f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|^2}\right)^T \left(\frac{\mathbf{x}}{\|\mathbf{x}\|^2}\right) \\ &= \|\mathbf{x}\|^2 \left(\lambda - f(\mathbf{z})^T \mathbf{z}\right) \geq 0 \end{aligned}$$

   since $\mathbf{z} = \mathbf{x} / \|\mathbf{x}\| \in S$.

3. Since f is symmetric

$$\begin{aligned} g(\mathbf{y},\mathbf{x}) &= \big(\lambda\mathbf{y} - f(\mathbf{y})\big)^T\mathbf{x} \\ &= \lambda\mathbf{y}^T\mathbf{x} - f(\mathbf{y})^T\mathbf{x} \\ &= \lambda\mathbf{x}^T\mathbf{y} - f(\mathbf{x})^T y \\ &= \big(\lambda\mathbf{x} - f(\mathbf{x})\big)^T\mathbf{y} = g(\mathbf{x},\mathbf{y}) \end{aligned}$$

4. g satisfies the conditions of Exercise 3.59 and therefore

$$(g(\mathbf{x},\mathbf{y}))^2 \leq g(\mathbf{x},\mathbf{x})g(\mathbf{y},\mathbf{y}) \text{ for every } \mathbf{x},\mathbf{y} \in X \tag{3.4}$$

By definition $g(\mathbf{x}_0,\mathbf{x}_0) = 0$ and (3.4) implies that

$$g(\mathbf{x}_0,\mathbf{y}) = 0 \text{ for every } \mathbf{y} \in X$$

That is

$$g(\mathbf{x}_0,\mathbf{y}) = \big(\lambda\mathbf{x}_0 - f(\mathbf{x}_0)\big)^T\mathbf{y} = 0 \text{ for every } y \in X$$

and therefore

$$\lambda\mathbf{x}_0 - f(\mathbf{x}_0) = \mathbf{0}$$

or

$$f(\mathbf{x}_0) = \lambda\mathbf{x}_0$$

In other words, $\mathbf{x}_0$ is an eigenvector. By construction, $\|\mathbf{x}_0\| = 1$.

3.88 1. Suppose $\mathbf{x}_2,\mathbf{x}_3 \in S$. Then

$$\big(\alpha\mathbf{x}_2 + \beta\mathbf{x}_3\big)^T\mathbf{x}_1 = \alpha\mathbf{x}_2^T\mathbf{x}_1 + \beta\mathbf{x}_3^T\mathbf{x}_1 = 0$$

so that $\alpha\mathbf{x}_2 + \beta\mathbf{x}_3 \in S$. S is a subspace.

Let $\{\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_n\}$ be a basis for X (Exercise 1.142). For $\mathbf{x} \in X$, there exists (Exercise 1.137) unique $\alpha_1,\alpha_2,\dots,\alpha_n$ such that

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \dots + \alpha_n\mathbf{x}_n$$

If $\mathbf{x} \in S$

$$\mathbf{x}^T\mathbf{x}_1 = \alpha_1\mathbf{x}_1^T\mathbf{x}_1 = 0$$

which implies that $\alpha_1 = 0$. Therefore, $\mathbf{x}_2,\mathbf{x}_3,\dots,\mathbf{x}_n$ span S and therefore $\dim S = n-1$.

2. For every $\mathbf{x} \in S$,

$$f(\mathbf{x})^T\mathbf{x}_0 = \mathbf{x}^T f(\mathbf{x}_0) = \mathbf{x}^T\lambda\mathbf{x}_0 = \lambda\mathbf{x}^T\mathbf{x}_0 = 0$$

since f is symmetric. Therefore $f(\mathbf{x}) \in \{\mathbf{x}_0\}^\perp = S$.

3.89 Let f be a symmetric operator. By the Spectral theorem (Proposition 3.6), there exists a diagonal matrix A which represents f. The elements of A are the eigenvalues of f. By Proposition 3.5, the determinant of A is the product of these diagonal elements.

3.90 By linearity

$$f(\mathbf{x}) = \sum_j x_j f(\mathbf{x}_j)$$

Q defines a quadratic form since

$$Q(\mathbf{x}) = \mathbf{x}^T f(\mathbf{x}) = \left(\sum_i x_i \mathbf{x}_i\right)^T \left(\sum_j x_j f(\mathbf{x}_j)\right) = \sum_i \sum_j x_i x_j \mathbf{x}_i^T f(\mathbf{x}_j) = \sum_i \sum_j a_{ij} x_i x_j$$

by Exercise 3.69.

3.91 Let f be the symmetric linear operator defining Q

$$Q(\mathbf{x}) = \mathbf{x}^T f(\mathbf{x})$$

By the Spectral theorem (Proposition 3.6), there exists an orthonormal basis $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ comprising the eigenvectors of f. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues, that is

$$f(\mathbf{x}_i) = \lambda_i \mathbf{x}_i \qquad i = 1, 2 \dots, n$$

Then for $\mathbf{x} = x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n$

$$\begin{aligned}
Q(\mathbf{x}) &= \mathbf{x}^T f(\mathbf{x}) \\
&= (x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n)^T f(x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n) \\
&= (x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n)^T (x_1 f(\mathbf{x}_1) + x_2 f(\mathbf{x}_2) + \cdots + x_n f(\mathbf{x}_n)) \\
&= (x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n)^T (x_1\lambda_1\mathbf{x}_1 + x_2\lambda_2\mathbf{x}_2 + \cdots + x_n\lambda_n\mathbf{x}_n) \\
&= x_1\lambda_1 x_1 + x_2\lambda_2 x_2 + \cdots + x_n\lambda_n x_n \\
&= \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2
\end{aligned}$$

3.92

1. Assuming that $a_{11} \neq 0$, the quadratic form can be rewritten as follows

$$\begin{aligned}
Q(x_1, x_2) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \\
&= a_{11}x_1^2 + 2a_{12}x_1x_2 + \frac{a_{12}^2}{a_{11}}x_2^2 - \frac{a_{12}^2}{a_{11}}x_2^2 + a_{22}x_2^2 \\
&= a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \left(\frac{a_{12}}{a_{11}}x_2\right)^2\right) + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 \\
&= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \left(\frac{a_{11}a_{22} - a_{12}^2}{a_{11}}\right)x_2^2
\end{aligned}$$

2. We observe that q must be positive for every x_1 and x_2 provided $a_{11} > 0$ and $a_{11}a_{22} - a_{12}^2 > 0$. Similarly q must be negative for every x_1 and x_2 if $a_{11} > 0$ and $a_{11}a_{22} - a_{12}^2 > 0$. Otherwise, we can choose values for x_1 and x_2 which make q both positive and negative.

   Note that the condition $a_{11}a_{22} > a_{12}^2 > 0$ implies that a_{11} and a_{12} must have the same sign.

3. If $a_{11} = a_{22} = 0$, then q is indefinite. Otherwise, if $a_{11} = 0$ but $a_{22} \neq 0$, then the q can we can "complete the square" using a_{22} and deduce

$$q \text{ is } \begin{Bmatrix} \text{positive} \\ \text{negative} \end{Bmatrix} \text{ semidefinite if and only if } \begin{Bmatrix} a_{11}, a_{22} \geq 0 \\ a_{11}, a_{22} \leq 0 \end{Bmatrix} \text{ and } a_{11}a_{22} \geq a_{12}^2$$

3.93 Let $Q\colon X \to \Re$ be a quadratic form on X. Then there exists a linear operator f such that

$$Q(\mathbf{x}) = \mathbf{x}^T f(\mathbf{x})$$

and (Exercise 3.13)

$$Q(\mathbf{0}) = \mathbf{0}^T f(\mathbf{0}) = 0$$

3.94 Suppose to the contrary that the positive (negative) definite matrix A is singular. Then there exists $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$ and therefore $\mathbf{x}'A\mathbf{x} = 0$ contradicting the definiteness of A.

3.95 Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ be the standard basis for $\Re^n$ (Example 1.79). Then for every i

$$\mathbf{e}_i' A \mathbf{e}_i = a_{ii} > 0$$

3.96 Let Q be the quadratic form defined by A. By Exercise 3.91, there exists an orthonormal basis such that

$$Q(\mathbf{x}) = \lambda_1 \mathbf{x}_1^2 + \lambda_2 \mathbf{x}_2^2 + \dots + \lambda_n \mathbf{x}_n^2$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A. This implies

$$\left\{\begin{matrix} Q(\mathbf{x}) > 0 \\ Q(\mathbf{x}) \geq 0 \\ Q(\mathbf{x}) < 0 \\ Q(\mathbf{x}) \leq 0 \end{matrix}\right\} \iff \left\{\begin{matrix} \lambda_i > 0 \\ \lambda_i \geq 0 \\ \lambda_i < 0 \\ \lambda_i \leq 0 \end{matrix}\right\} \qquad i = 1, 2, \dots, n$$

3.97 Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A. By Exercise 3.89

$$\det(A) = \lambda_1 \lambda_2 \dots \lambda_n$$

By Exercise 3.96, $\lambda_i \geq 0$ for every i and therefore $\det(A) \geq 0$. We conclude that

$$\det(A) > 0 \iff \lambda_i > 0 \text{ for every } i \iff A \text{ is positive definite}$$

by Exercise 3.96.

3.98 1. $A\mathbf{0} = \mathbf{0}$. Therefore, $\mathbf{0}$ is always a solution.

2. Assume $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions, that is

$$A\mathbf{x}_1 = \mathbf{0} \text{ and } A\mathbf{x}_2 = \mathbf{0}$$

Then

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0}$$

$\mathbf{x}_1 + \mathbf{x}_2$ is also a solution.

3. Let f be the linear function defined by

$$f(\mathbf{x}) = A\mathbf{x}$$

The system of equations $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if

$$\text{kernel } f \neq \{\mathbf{0}\} \iff \text{nullity } f > 0$$

By the Rank theorem (Exercise 3.24)

$$\text{rank} f + \text{nullity} f = \dim X$$

so that

$$\text{nullity } f > 0 \iff \text{rank} f < \dim X = n$$

3.99 1. Assume $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions of (3.16). That is

$$A\mathbf{x}_1 = \mathbf{c} \text{ and } A\mathbf{x}_2 = \mathbf{c}$$

Subtracting

$$A\mathbf{x}_1 - A\mathbf{x}_2 = A(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$$

2. Assume $\mathbf{x}_p$ solves (3.16) while $\mathbf{x}$ is any solution to (3.17). That is

$$A\mathbf{x}_p = \mathbf{c} \text{ and } A\mathbf{x} = \mathbf{0}$$

Adding

$$A\mathbf{x}_p + A\mathbf{x} = A(\mathbf{x}_p + \mathbf{x}) = \mathbf{c}$$

We conclude that $\mathbf{x}_p + \mathbf{x}$ solves (3.16) for every $\mathbf{x} \in K$.

3. If $\mathbf{0}$ is the only solution of (3.17), $K = \{\mathbf{0}\}$. Assume $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions of (3.16). Then $\mathbf{x}_1 - \mathbf{x}_2 \in K = \{0\}$ which implies $\mathbf{x}_1 = \mathbf{x}_2$.

3.100 Let $S = \{\, \mathbf{x} : A\mathbf{x} = c \,\}$. For every $\mathbf{x}, \mathbf{y} \in S$ and $\alpha \in \Re$

$$A\alpha\mathbf{x} + (1-\alpha)\mathbf{y} = \alpha A\mathbf{x} + (1-\alpha)A\mathbf{y} = \alpha\mathbf{c} + (1-\alpha)\mathbf{c} = c$$

Therefore, $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in S$. S is affine.

3.101 Let $S \neq \emptyset$ be an affine set $\Re^n$. Then there exists a unique subspace V such that

$$S = \mathbf{x}_0 + V$$

for some $\mathbf{x}_0 \in S$ (Exercise 1.150). The orthogonal complement of V is

$$V^\perp = \{\, \mathbf{a} \in X : \mathbf{a}\mathbf{x} = 0 \text{ for every } \mathbf{x} \in V \,\}$$

Let $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ be a basis for $V^\perp$. Then

$$V = (V^\perp)^\perp = \{\mathbf{x} : \mathbf{a}_i\mathbf{x} = 0, \qquad i = 1, 2, \dots m\}$$

Let A be the $m \times n$ matrix whose rows are $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$. Then V is the set of solutions to the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, that is

$$V = \{\, x : A\mathbf{x} = \mathbf{0} \,\}$$

Therefore

$$\begin{aligned} S &= \mathbf{x}_0 + V \\ &= \mathbf{x}_0 + \{\, \mathbf{x} : A\mathbf{x} = \mathbf{0} \,\} \\ &= \{\, \mathbf{x} : A(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \,\} \\ &= \{\, \mathbf{x} : A\mathbf{x} = \mathbf{c} \,\} \end{aligned}$$

where $\mathbf{c} = A\mathbf{x}_0$.

3.102 Consider corresponding homogeneous system

$$\begin{aligned} x_1 + 3x_2 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$$

Multiplying the second equation by 3

$$\begin{aligned} x_1 + 6x_2 &= 0 \\ 3x_1 - 3x_2 &= 0 \end{aligned}$$

and adding yields

$$4x_1 = 0$$

for which the only solution is $x_1 = 0$. Substituting in the first equation implies $x_2 = 0$. The kernel of $f = A\mathbf{x}$ is $\{0\}$. Therefore $\dim f(\Re^2) = 2$, and the system $A\mathbf{x} = c$ has a unique solution for every c_1, c_2.

3.103 We can write the system $A\mathbf{x} = \mathbf{c}$ in the form

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Subtracting $\mathbf{c}$ from the jth column gives

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + \begin{pmatrix} x_j a_{1j} - c_1 \\ \vdots \\ x_j a_{nj} - c_n \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \mathbf{0}$$

so that the columns of the matrix

$$C = \begin{pmatrix} a_{11} & \dots & (x_j a_{1j} - c_1) & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & (x_j a_{nj} - c_n) & \dots & a_{nn} \end{pmatrix}$$

are linearly dependent (Exercise 1.133). Therefore $\det(C) = 0$. Let B_j denote the matrix obtained from A by replacing the jth column with $\mathbf{c}$. Then A, B_j and C differ only in the jth column, with the jth column of C being a linear combination of the jth columns of A and B_j.

$$\begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix} = x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} - \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix}$$

By Exercise 3.81

$$\det(C) = \mathbf{x}_j \det(A) - \det(B_j) = 0$$

and therefore

$$x_j = \frac{\det(B_j)}{\det(A)}$$

as required.

3.104 Let

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

The inverse satisfies the equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In particular, this means that A and C satisfy the equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

By Cramer's rule (Exercise 3.103

$$A = \frac{\begin{vmatrix} 1 & b \\ 0 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{d}{\Delta}$$

where $\Delta = ad - bc$. Similarly

$$C = \frac{\begin{vmatrix} a & 1 \\ c & 0 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{-c}{\Delta}$$

B and D are determined analogously.

3.105 A portfolio is duplicable if and only if there is a different portfolio $\mathbf{y} \neq \mathbf{x}$ such that

$$R\mathbf{x} = R\mathbf{y}$$

or

$$R(\mathbf{x} - \mathbf{y}) = \mathbf{0}$$

There exists a duplicable portfolio if and only if this homogeneous system has a non-trivial solution, that is if rank $R < A$.

3.106 State $\bar{s}$ is insurable if there is a solution to the linear system

$$R\mathbf{x} = \mathbf{e}_{\bar{s}} \tag{3.5}$$

where $\mathbf{e}_{\bar{s}}$ is the $\bar{s}$-th unit vector (the $\bar{s}$ Arrow-Debreu security. (3.5) has a solution for every state s if and only if $f(\Re^A) = \Re^S$, that is rank $R = S$.

3.107 Figure 3.1.

3.108 Let S be an affine subset of $\Re^n$. Then there exists (Exercise 3.101) a system of linear equations $A\mathbf{x} = \mathbf{c}$ such that

$$S = \{\, \mathbf{x} : A\mathbf{x} = \mathbf{c} \,\}$$

Let $\mathbf{a}_i$ denote the i-th row of A. Then

$$\begin{aligned} S &= \{\, x : \mathbf{a}_i\mathbf{x} = c_i, \qquad i = 1, 2, \dots, n \,\} \\ &= \bigcap_{i=1}^{n} \{\, \mathbf{x} : \mathbf{a}_i\mathbf{x} = c_i \,\} \end{aligned}$$

where each $\{\, \mathbf{x} : \mathbf{a}_i\mathbf{x} = c_i \,\}$ is a hyperplane in $\Re^n$ (Example 3.21).

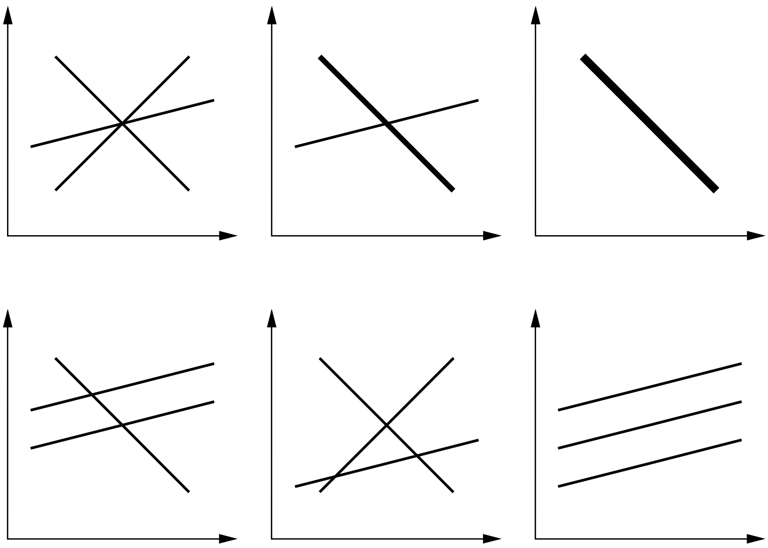

Figure 3.1: The solutions of three equations in two unknowns

3.109 Let $S = \{ \mathbf{x} : A\mathbf{x} \leq \mathbf{c} \}$. For every $\mathbf{x}, \mathbf{y} \in S$ and $0 \leq \alpha \leq 1$

$$A\mathbf{x} \leq \mathbf{c}$$
$$A\mathbf{y} \leq \mathbf{c}$$

and therefore

$$A\alpha\mathbf{x} + (1-\alpha)\mathbf{y} = \alpha A\mathbf{x} + (1-\alpha)A\mathbf{y} \leq \alpha\mathbf{c} + (1-\alpha)\mathbf{c} = c$$

Therefore, $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in S$. S is a convex set.

3.110 We have already seen that $S = \{ \mathbf{x} : A\mathbf{x} \leq \mathbf{0} \}$ is convex. To show that it is a cone, let $\mathbf{x} \in S$. Then

$$A\mathbf{x} \leq \mathbf{0}$$
$$A\alpha\mathbf{x} \leq \mathbf{0}$$

so that $\alpha\mathbf{x} \in S$. S is a convex cone.

3.111

1. Each column A_j is a vector in $\Re^m$. If the set $\{A_1, A_2, \dots, A_k\}$ is linearly independent, it has at most m elements, that is $k \leq m$ and $\mathbf{x}$ is a basic feasible solution.
2. (a) Assume $\{A_1, A_2, \dots, A_k\}$ are linearly dependent. Then (Exercise 1.133) there exist numbers $y_1, y_2, \dots, y_k$, not all zero, such that

   $$y_1 A_1 + y_2 A_2 + \cdots + y_k A_k = \mathbf{0}$$

   $\mathbf{y} = (y_1, y_2, \dots, y_k)$ is a nontrivial solution to the homogeneous system.

   (b) For every $t \in \Re$, $-t\mathbf{y} \in$ kernel $f = A\mathbf{x}$ and $\mathbf{x}' = \mathbf{x} - t\mathbf{y}$ is a solution of the corresponding nonhomogeneous system $A\mathbf{x} = \mathbf{c}$. To see this directly, subtract

   $$At\mathbf{y} = \mathbf{0}$$

from

$$A\mathbf{x} = \mathbf{c}$$

to give

$$A\mathbf{x}' = A(\mathbf{x} - t\mathbf{y}) = \mathbf{c}$$

(c) Note that $\mathbf{x} > 0$ and therefore $\hat{t} > 0$ which implies that $\hat{x_j} > 0$ for every $y_j \leq 0$. For every $y_j > 0$, $x_j/y_j \geq \hat{t}$, which implies that $x_j \geq \hat{t}y_j$, so that

$$\hat{x_j} \geq x_j - \hat{t}y_j \geq 0$$

Therefore, $\hat{\mathbf{x}}$ is a feasible solution.

(d) There exists some coordinate h such that $\hat{t} = x_h/y_h$ so that

$$\hat{x}_h = x_h - \hat{t}y_h = 0$$

so that

$$\mathbf{c} = \sum_{\substack{j\ =1 \\ j \neq h}}^{k} \hat{x_j} A_J$$

$\hat{x}$ is a feasible solution with one less positive component.

3. Starting with any nonbasic feasible solution, this elimination technique can be repeated until the remaining vectors are linearly independent and a basic feasible solution is obtained.

3.112 1. Exercise 1.173.

2. For each i, there exists l_i elements $\mathbf{x}_{ij}$ and coefficients $a_{ij} > 0$ such that

$$\mathbf{x}_i = \sum_{i=1}^{l_i} a_{ij}\mathbf{x}_{ij}$$

and $\sum_{j=1}^{l_i} a_{ij} = 1$. Hence

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{x}_i = \sum_{i=1}^{n} \sum_{j=1}^{l_i} a_{ij}\mathbf{x}_{ij}$$

3. Direct computation.

4. Regarding the a_{ij} as "variables" and the points z_{ij} as coefficents,

$$\mathbf{z} = \sum_{i=1}^{n} \sum_{j=1}^{l_i} a_{ij}\mathbf{z}_{ij}$$

is a linear equation system in which variables are restricted to be nonnegative. By the fundamental theorem of linear programming (Exercise 3.111), there exists a basic feasible solution. That is, there exists coefficients $b_{ij} \geq 0$ and $b_{ij} > 0$ for at most $(m+n)$ components such that

$$\mathbf{z} = \sum_{i=1}^{n} \sum_{j=1}^{l_i} b_{ij} \mathbf{z}_{ij} \tag{3.6}$$

Decomposing, (3.6) implies

$$\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{l_i} b_{ij} \mathbf{x}_{ij}$$

and

$$\sum_{j=1}^{l_i} b_{ij} = 1 \qquad \text{for every } \ i$$

5. (3.6) implies that at least one $b_{ij} > 0$ for every i. This accounts for at least n of the positive b_{ij}. Since there are at most $(m+n)$ coefficients b_{ij} which are strictly positive, there are at most m indices i which have more than one positive coefficient b_{ij}. For the remaining $m-n$ indices, $\mathbf{x}_i = \mathbf{x}_{ij}$ for some j; that is $\mathbf{x}_i \in S_i$.

3.113 1. Since A is productive, there exists $\mathbf{x} \geq 0$ such that $A\mathbf{x} > 0$. Consider any $\mathbf{z}$ for which $A\mathbf{z} \geq \mathbf{0}$. For every $\alpha > 0$

$$A(\mathbf{x} + \alpha \mathbf{z}) = A\mathbf{x} + \alpha A \mathbf{z} > 0 \tag{3.7}$$

Suppose to the contrary that $\mathbf{z} \ngeq \mathbf{0}$. That is, there exists some component $z_i < 0$. Let

$$\alpha = \max\{-\frac{z_i}{x_i}\}$$

Without loss of generality, z_1 attains this maximum, that is assume $\alpha = z_1/x_1$. Then

$$x_1 + \alpha z_1 = 0$$

and

$$x_i + \alpha z_i \geq 0$$

for every i.

Now consider the matrix $B = I - A$. By the assumptions of the Leontief model (Example 3.35), the matrix A has 1 along the diagonal and negative off-diagonal elements. That is

$$\begin{aligned} a_{ii} &= 1 \qquad i = 1, 2, \ldots, n \\ a_{ij} &\leq 0 \qquad i, j = 1, 2, \ldots, n, \quad j \neq j \end{aligned}$$

Therefore

$$B = I - A \geq \mathbf{0}$$

That is, every element of B is nonnegative. Consequently since $\mathbf{x} + \alpha \mathbf{z} \geq \mathbf{0}$

$$B(\mathbf{x} + \alpha \mathbf{z}) \geq \mathbf{0} \tag{3.8}$$

On the other hand, substituting $A = I - B$ in (3.8)

$$(I - B)(\mathbf{x} + \alpha \mathbf{z}) > \mathbf{0}$$
$$\mathbf{x} + \alpha \mathbf{z} > B(\mathbf{x} + \alpha \mathbf{z})$$

which implies that the first component of $B(\mathbf{x} + \alpha \mathbf{z})$ is negative, contradicting (3.8). This contradiction establishes that $\mathbf{z} \geq \mathbf{0}$.

Suppose $A\mathbf{x} = \mathbf{0}$. *A fortiori* $A\mathbf{x} \geq \mathbf{0}$. By the previous part this implies $\mathbf{x} \geq \mathbf{0}$. On the other hand, it also implies that $-A\mathbf{x} = A(-\mathbf{x}) = \mathbf{0}$ so that $-\mathbf{x} \geq \mathbf{0}$. We conclude that $\mathbf{x} = \mathbf{0}$ is the only solution to $A\mathbf{x} = \mathbf{0}$. A is nonsingular.

Since A is nonsingular, the system $A\mathbf{x} = \mathbf{y}$ has a unique solution $\mathbf{x}$ for any $\mathbf{y} \geq \mathbf{0}$. By the first part, $\mathbf{x} \geq \mathbf{0}$.

3.114 Suppose A is productive. By the previous exercise, A is nonsingular with inverse A^{-1}. Let $\mathbf{e}_i$ be the ith unit vector. Since $\mathbf{e}_i \geq \mathbf{0}$, there exists $\mathbf{x}_i \geq \mathbf{0}$ such that

$$A\mathbf{x}_i = \mathbf{e}_i$$

Multiplying by A^{-1}

$$\mathbf{x}_i = A^{-1}A\mathbf{x}_i = A^{-1}\mathbf{e}_i = A_i^{-1}$$

where A_i^{-1} is the i column of A^{-1}. Since $\mathbf{x}_i \geq \mathbf{0}$ for every i, we conclude that $A^{-1} \geq \mathbf{0}$.

Conversely, assume that $A^{-1} \geq \mathbf{0}$. Let $\mathbf{1} = (1, 1, \ldots, 1)$ denote a net output of 1 for each commodity. Then

$$\mathbf{x} = A^{-1}\mathbf{1} \geq \mathbf{0}$$

and

$$A\mathbf{x} = \mathbf{1} > \mathbf{0}$$

A is productive.

3.115 Takayama 1985, p.383, Theorem 4.C.4.

3.116 Let $\mathbf{a}_0 = (a_{01}, a_{02}, \ldots, a_{0n})$ be the vector of labour requirements and w the wage rate. The unit profit of industry i is

$$\pi_i = p_i + \sum_{j \neq i} a_{ij} p_j - w a_0$$

Recall that $a_{ij} \leq 0$ for $j \neq i$. The vector of unit profits for all industries is

$$\Pi = A\mathbf{p} - w a_0$$

Profits will be zero in all industries if there exists a price system $\mathbf{p}$ such that

$$\Pi = A\mathbf{p} - w a_0 = \mathbf{0}$$

or

$$A\mathbf{p} = w a_0 \tag{3.9}$$

By the previous results, (3.9) has a unique nonnegative solution $\mathbf{p} = A^{-1} w a_0$ if the technology A is productive. Furthermore, A^{-1} is nonnegative. Since $a_0 > 0$, so is $\mathbf{p} > 0$.

3.117 Let u_B denote the steady state unemployment rate for blacks. Then u_B satisfies the equation

$$u_B = 0.0038(1 - u_B) + 0.8975u_B$$

which implies that $u_B = 0.036$. That is, the data implies an unemployment rate of 3.6 percent for blacks. Similarly, the unemployment rate for white males u_W satisfies the equation

$$u_W = 0.0022(1 - u_W) + 0.8614u_W$$

which implies that $u_W = 0.016$ or 1.6 percent.

3.118 The transition matrix is

$$T = \begin{pmatrix} .6 & .25 \\ .4 & .75 \end{pmatrix}$$

If the current state vector is $\mathbf{x}^0 = (.4, .6)$, the state vector after a single mailing will be

$$\begin{aligned} \mathbf{x}^1 &= T\mathbf{x}^0 \\ &= \begin{pmatrix} .6 & .25 \\ .4 & .75 \end{pmatrix} \begin{pmatrix} .4 \\ .6 \end{pmatrix} \\ &= \begin{pmatrix} 0.39 \\ .61 \end{pmatrix} \end{aligned}$$

Following a single mailing, the number of subscribers will drop to 30 percent of the mailing list, comprising 24 percent from renewals and 15 percent new subscriptions.

3.119 Let $f(x) = x^2$. For every $x_1, x_2 \in \Re$ and $0 \le \alpha \le 1$

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= (\alpha x_1 + (1-\alpha)x_2)^2 \\ &= (\alpha x_1 + (1-\alpha)x_2)(\alpha x_1 + (1-\alpha)x_2) \\ &= \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1x_2 + (1-\alpha)^2 x_2^2 \\ &= \alpha x_1^2 + (1-\alpha)x_2^2 - \alpha x_1^2 - (1-\alpha)x_2^2 + \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1x_2 + (1-\alpha)^2 x_2^2 \\ &= \alpha x_1^2 + (1-\alpha)x_2^2 - \Big(\alpha(1-\alpha)x_1^2 - 2\alpha(1-\alpha)x_1x_2 + \alpha(1-\alpha)x_2^2\Big) \\ &= \alpha x_1^2 + (1-\alpha)x_2^2 - \alpha(1-\alpha)(x_1 - x_2)^2 \\ &\le \alpha x_1^2 + (1-\alpha)x_2^2 \\ &= \alpha f(x_1) + (1-\alpha)(x_2) \end{aligned}$$

3.120 $f(x) = x$ is linear and therefore convex. In the previous exercise we showed that x^2 is convex. Therefore $f(x) = x^n$ is convex for $n = 1, 2$. Assume that f is convex for

$n-1$. Then

$$\begin{aligned}
f(\alpha x_1 + (1-\alpha)x_2) &= (\alpha x_1 + (1-\alpha)x_2)^n \\
&= (\alpha x_1 + (1-\alpha)x_2)(\alpha x_1 + (1-\alpha)x_2)^{n-1} \\
&\leq (\alpha x_1 + (1-\alpha)x_2)(\alpha x_1^{n-1} + (1-\alpha)x_2^{n-1}) \qquad \text{(since } x^{n-1} \text{ is convex)} \\
&= \alpha^2 x_1^n + \alpha(1-\alpha)x_1^{n-1}x_2 + \alpha(1-\alpha)x_1 x_2^{n-1} + (1-\alpha)^2 x_2^n \\
&= \alpha x_1^n + (1-\alpha)x_2^n - \alpha x_1^n - (1-\alpha)x_2^n \\
&\qquad + \alpha^2 x_1^n + \alpha(1-\alpha)x_1^{n-1}x_2 + \alpha(1-\alpha)x_1 x_2^{n-1} + (1-\alpha)^2 x_2^n \\
&= \alpha x_1^n + (1-\alpha)x_2^n - \alpha(1-\alpha)\Big(x_1^n - x_1 x_2^{n-1} - x_1^{n-1}x_2 + x_2^n\Big) \\
&= \alpha x_1^n + (1-\alpha)x_2^n - \alpha(1-\alpha)\Big(x_1^{n-1}(x_1 - x_2) - x_2^{n-1}(x_1 - x_2)\Big) \\
&= \alpha x_1^n + (1-\alpha)x_2^n - \alpha(1-\alpha)\Big((x_1 - x_2)(x_1^{n-1} - x_2^{n-1})\Big)
\end{aligned}$$

Since x^m is monotonic (Example 2.53)

$$x_1^{n-1} - x_2^{n-1} \geq 0 \iff x_1 - x_2 \geq 0$$

and therefore

$$(x_1 - x_2)(x_1^{n-1} - x_2^{n-1}) \geq 0$$

We conclude that

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha x_1^n + (1-\alpha)x_2^n = \alpha f(x_1) + (1-\alpha)(x_2)$$

f is convex for all $n = 1, 2, \dots$.

3.121 For given $\mathbf{x}_1, \mathbf{x}_2 \in S$, define $g\colon [0,1] \to S$ by

$$g(t) = (1-t)\mathbf{x}_1 + t\mathbf{x}_2$$

Then $g(0) = \mathbf{x}_1$, $g(1) = \mathbf{x}_2$ and $h = g \circ f$.

Assume f is convex. For any $t_1, t_2 \in [0,1]$, let

$$g(t_1) = \bar{\mathbf{x}}_1 \text{ and } \; g(t_2) = \bar{\mathbf{x}}_2$$

For any $\alpha \in [0,1]$

$$\begin{aligned}
g\big(\alpha t_1 + (1-\alpha)t_2\big) &= \alpha\bar{\mathbf{x}}_1 + (1-\alpha)\bar{\mathbf{x}}_2 \\
h\big(\alpha t_1 + (1-\alpha)t_2\big) &= f\big(\alpha\bar{\mathbf{x}}_1 + (1-\alpha)\bar{\mathbf{x}}_2\big) \\
&\leq \alpha f(\bar{\mathbf{x}}_1) + (1-\alpha)f(\bar{\mathbf{x}}_2) \\
&\leq \alpha h(t_1) + (1-\alpha)t_2)
\end{aligned}$$

h is convex.

Conversely, assume h is convex for any $\mathbf{x}_1, \mathbf{x}_2 \in S$. For any $\alpha \in [0,1]$

$$g(\alpha) = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$$

and

$$\begin{aligned}
f\big(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2\big) &= h(\alpha) \\
&\leq \alpha h(0) + (1-\alpha)h(1) \\
&= \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)
\end{aligned}$$

Since this is true for any $\mathbf{x}_1, \mathbf{x}_2 \in S$, we conclude that f is convex.

3.122 Assume f is convex which implies epi f is convex. The points $(\mathbf{x}_i, f(\mathbf{x}_i)) \in \text{epi } f$. Since epi f is convex

$$\alpha_1(\mathbf{x}_1, f(\mathbf{x}_1)) + \alpha_2(\mathbf{x}_1, f(\mathbf{x}_1)) + \cdots + (\mathbf{x}_n, f(\mathbf{x}_n)) \in \text{epi } f$$

that is

$$f(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n) \leq \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_1) + \cdots + \alpha_n f(\mathbf{x}_n))$$

Conversely, letting $n = 2$ and $\alpha = \alpha_1$, (3.25) implies that

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)$$

Jensen's inequality can also be proved by induction from the definition of a convex function (see for example Sydsaeter + Hammond 1995; p.624).

3.123 For each i, let $y_i = \log x_i$ so that

$$x_i = e^{y_i}$$
$$x_i^{\alpha_i} = e^{\alpha_i y_i}$$

Since e^x is convex (Example 3.41)

$$x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n} \qquad a_i > 0 = \prod \exp(\alpha_i y_i) = \exp\left(\sum \alpha_i y_i\right) \leq \sum \alpha_i e^{y_i} = \sum \alpha_i x_i$$

by Jensen's inequality. Setting $\alpha_i = 1/n$, we have

$$(x_1 x_2 \ldots x_n)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} x_i$$

as required.

3.124 Assume f is concave. That is for every $\mathbf{x}_1, \mathbf{x}_2 \in S$ and $0 \leq \alpha \leq 1$

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)$$

Multiplying through by -1 reverses the inequality so that

$$-f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq -\alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) = \alpha - f(\mathbf{x}_1) + (1-\alpha) - f(\mathbf{x}_2)$$

which shows that $-f$ is concave. The converse follows analogously.

3.125 Assume that f is concave. Then $-f$ is convex and by Theorem 3.7

$$\text{epi } -f = \{\, (x,y) \in X \times \Re : y \geq -f(x), x \in X \,\}$$

is convex. But

$$\text{epi } -f = \{\, (x,y) \in X \times \Re : y \geq -f(x), x \in X \,\} = \{\, (x,y) \in X \times \Re : y \leq f(x), x \in X \,\} = \text{hypo } f$$

Therefore hypo f is convex.

Conversely, if hypo f is convex, epi $-f$ is convex which implies that $-f$ is convex and hence f is concave.

3.126 Suppose that $\mathbf{x}^1$ minimizes the cost of producing y at input prices $\mathbf{w}^1$ while $\mathbf{x}^2$ minimizes cost at $\mathbf{w}^2$. For some $\alpha \in [0,1]$, let $\bar{\mathbf{w}}$ be the weighted average price, that is

$$\bar{\mathbf{w}} = \alpha \mathbf{w}^1 + (1-\alpha)\mathbf{w}^2$$

and suppose that $\bar{\mathbf{x}}$ minimizes cost at $\bar{\mathbf{w}}$. Then

$$\begin{aligned} c(\bar{\mathbf{w}}, y) &= \bar{\mathbf{w}}\bar{\mathbf{x}} \\ &= (\alpha \mathbf{w}^1 + (1-\alpha)\mathbf{w}^2)\bar{\mathbf{x}} \\ &= \alpha \mathbf{w}^1 \bar{\mathbf{x}} + (1-\alpha)\mathbf{w}^2\bar{\mathbf{x}} \end{aligned}$$

But since $\mathbf{x}^1$ and $\mathbf{x}^2$ minimize cost at $\mathbf{w}^1$ and $\mathbf{w}^2$ respectively

$$\begin{aligned} \alpha \mathbf{w}^1 \bar{\mathbf{x}} &\geq \alpha \mathbf{w}^1 \mathbf{x}^1 = \alpha c(\mathbf{w}^1, y) \\ (1-\alpha)\mathbf{w}^2\bar{\mathbf{x}} &\geq (1-\alpha)\mathbf{w}^2\mathbf{x}^2 = (1-\alpha)c(\mathbf{w}^2, y) \end{aligned}$$

so that

$$c(\bar{\mathbf{w}}, y) = c(\alpha \mathbf{w}^1 + (1-\alpha)\mathbf{w}^2, y) = \alpha \mathbf{w}^1\bar{\mathbf{x}} + (1-\alpha)\mathbf{w}^2\bar{\mathbf{x}} \geq \alpha c(\mathbf{w}^1, y) + (1-\alpha)c\mathbf{w}^2, y)$$

This establishes that the cost function c is concave in $\mathbf{w}$.

3.127 Since u is concave, Jensen's inequality implies

$$u\left(\sum_{t=1}^{T} \frac{1}{T} c_t\right) \geq \sum_{t=1}^{T} \frac{1}{T} u(c_t) = \frac{1}{T}\sum_{t=1}^{T} u(c_t)$$

for any consumption stream $c_1, c_2, \ldots, c_T$ so that

$$U = \sum_{t=1}^{T} u(c_t) \leq Tu\left(\sum_{t=1}^{T} \frac{1}{T} c_t\right) = Tu(\bar{c})$$

It is impossible to do better than consume a constant fraction $\bar{c} = w/T$ of wealth in each period.

3.128 If $x_1 = x_3$, the inequality is trivially satisfied. Now assume $x_1 \neq x_3$. Since $x_2 \in [x_1, x_3]$, there exists $\alpha \in [0,1]$ such that

$$x_2 = \alpha x_1 + (1-\alpha)x_2$$

Let $\bar{x} = x_1 - x_2 + x_3$. Then $\bar{x} \in [x_1, x_3]$ and there exists $\beta \in [0,1]$ such that

$$\bar{x} = \beta x_1 + (1-\beta)x_2$$

Adding

$$\bar{x} + x_2 = (\alpha + \beta)x_1 + \big((1-\alpha) + (1-\beta)\big)x_3$$

or

$$x_1 - x_3 = (\alpha + \beta)(x_3 - x_1)$$

which implies that $\alpha + \beta = 1$ and therefore $\beta = 1 - \alpha$. Since f is convex

$$\begin{aligned} f(x_2) &\leq \alpha f(x_1) + (1-\alpha)f(x_2 \\ f(\bar{x}) &\leq \beta f(x_1) + (1-\beta)f(x_2) \\ &= (1-\alpha)f(x_1) + \alpha f(x_3) \end{aligned}$$

Adding

$$f(\bar{x}) + f(x_2) \leq f(x_1) + f(x_3)$$

3.129 Let $x_1, x_2, y_1, y_2 \in \Re$ with $x_1 < x_2$ and $y_1 < y_2$. Note that $x_1 - y_2 \leq x_2 - y_2 \leq x_2 - y_1$ and therefore (Exercise 3.128)

$$f\big(x_1 - y_2) - (x_2 - y_2) + (x_2 - y_1)\big) > f(x_1 - y_2) - f(x_2 - y_2) + f(x_2 - y_1)$$

That is

$$f(x_1 - y_1) > f(x_1 - y_2) - f(x_2 - y_2) + f(x_2 - y_1)$$

Rearranging

$$f(x_2 - y_2) - f(x_1 - y_2) > f(x_2 - y_1) - f(x_1 - y_1)$$

as required.

3.130 A functional is affine if and only if inequalities (3.24) and (3.26) are satisfied as equalities.

3.131 Since f and g are convex on S

$$f(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) \leq \beta f(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2) \tag{3.10}$$
$$g(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) \leq \beta g(\mathbf{x}^1) + (1-\beta) g(\mathbf{x}^2) \tag{3.11}$$

for every $\mathbf{x}^1, \mathbf{x}^2 \in S$ and $\beta \in [0,1]$. Adding

$$(f+g)(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) \leq \beta (f+g)(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2)$$

$f + g$ is convex. Multiplying (3.10) by $\alpha \geq 0$

$$\begin{aligned} \alpha f(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) &\leq \alpha(\beta f(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2)) \\ &= (\beta \alpha f(\mathbf{x}^1) + (1-\beta)\alpha f(\mathbf{x}^2)) \end{aligned}$$

αf is convex.

Moreover, if f is strictly convex,

$$f(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) < \beta f(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2) \tag{3.12}$$

for every $\mathbf{x}^1, \mathbf{x}^2 \in S$, $\mathbf{x}^1 \neq \mathbf{x}^2$ and $\beta \in (0,1)$. Adding this to (3.11)

$$(f+g)(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) < \beta (f+g)(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2)$$

so that $f + g$ is strictly convex. Multiplying (3.12) by $\alpha > 0$

$$\begin{aligned} \alpha f(\beta \mathbf{x}^1 + (1-\beta)\mathbf{x}^2) &< \alpha(\beta f(\mathbf{x}^1) + (1-\beta) f(\mathbf{x}^2)) \\ &= (\beta \alpha f(\mathbf{x}^1) + (1-\beta)\alpha f(\mathbf{x}^2)) \end{aligned}$$

αf is strictly convex.

3.132

$$\mathbf{x} \in \text{epi } (f \vee g) \iff \mathbf{x} \in \text{epi } f \text{ and } \mathbf{x} \in \text{epi } g$$

That is

$$\text{epi } (f \vee g) = \text{epi } f \cap \text{epi } g$$

Therefore epi $f \vee g$ is convex (Exercise 1.162) and therefore f is convex (Proposition 3.7).

3.133 If f is convex

$$f(\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2) \leq \alpha f(\mathbf{x}^1) + (1-\alpha)f(\mathbf{x}^2)$$

Since g is increasing

$$\begin{aligned} g\big(f(\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2)\big) &\leq g\big(\alpha f(\mathbf{x}^1) + (1-\alpha)f(\mathbf{x}^2)\big) \\ &\leq \alpha g\big(f(\mathbf{x}^1)\big) + (1-\alpha)g\big(f(\mathbf{x}^2)\big) \end{aligned}$$

since g is also convex. The concave case is proved similarly.

3.134 Let $F = \log f$. If F is convex, $f(\mathbf{x}) = e^{F(\mathbf{x})}$ is an increasing convex function of a convex function and is therefore convex (Exercise 3.133).

3.135 If f is positive and concave, then $\log f$ is concave (Exercise 3.51). Therefore

$$\log\frac{1}{f} = \log 1 - \log f = -\log f$$

is convex. By the previous exercise (Exercise 3.134), this implies that $1/f$ is convex.

If f is negative and convex, then $-f$ is positive and concave, $1/-f$ is convex, and therefore $1/f$ is concave.

3.136 Consider the identity

$$\begin{aligned} g\big(f(x_1\vee x_2)\big) &+ g\big(f(x_1\wedge x_2)\big) - g\big(f(x_1)\big) - g\big(f(x_2)\big) \\ &= g\big(f(x_1\vee x_2)\big) + g\big(f(x_1\wedge x_2)\big) - g\big(f(x_1)\big) - g\big(f(x_1\vee x_2)\big) + f(x_1\wedge x_2)\big) - f(x_1)\big) \\ &\quad + g\big(f(x_1\vee x_2) + f(x_1\wedge x_2) - f(x_1)\big) - g\big(f(x_2)\big) \end{aligned} \tag{3.13}$$

Define

$$\varphi(x_1,x_2) = g\big(f(x_1\vee x_2)\big) + g\big(f(x_1\wedge x_2)\big) - g\big(f(x_1)\big) - g\big(f(x_2)\big)$$

Then $g\circ f$ is supermodular if φ is nonnegative definite and submodular if φ is nonpositive definite. Using the identity (3.13), φ can be decomposed into two components

$$\begin{aligned} \varphi(x_1,x_2) &= \varphi_1(x_1,x_2) + \varphi_2(x_1,x_2) \\ \varphi_1(x_1,x_2) &= g\big(f(x_1\vee x_2)\big) + g\big(f(x_1\wedge x_2)\big) - g\big(f(x_1)\big) \\ &\qquad - g\big(f(x_1\vee x_2) + f(x_1\wedge x_2) - f(x_1)\big) \\ \varphi_2(x_1,x_2) &= g\big(f(x_1\vee x_2) + f(x_1\wedge x_2) - f(x_1)\big) - g\big(f(x_2)\big) \end{aligned} \tag{3.14}$$

φ will definite if both components are definite.

For any $x_1, x_2 \in x_1$, let $a = f(x_1\wedge x_2)$, $b = f(x_1)$ and $c = f(x_1\vee x_2)$. Provided f is monotone, b lies between a and c. Substituting in (3.14)

$$\varphi_1(x_1,x_2) = g(c) + g(a) - g(b) - g(c+a-b)$$

and Exercise 3.128 implies

$$\varphi_1(x_1,x_2) = g(c) + g(a) - g(b) - g(c+a-b) \begin{Bmatrix} \geq O \\ \leq 0 \end{Bmatrix} \text{ if } g \text{ is } \begin{Bmatrix} \text{convex} \\ \text{concave} \end{Bmatrix} \tag{3.15}$$

Now consider φ_2.

$$f(x_1\vee x_2) + f(x_1\wedge x_2) - f(x_1) \text{ is } \begin{Bmatrix} \geq \\ \leq \end{Bmatrix} f(x_2) \text{ if } f \text{ is } \begin{Bmatrix} \text{supermodular} \\ \text{submodular} \end{Bmatrix}$$

and therefore since g is increasing

$$\varphi_2(x_1, x_2) = \begin{cases} \geq 0 & \text{if } f \text{ is supermodular} \\ \leq 0 & \text{if } f \text{ is submodular} \end{cases} \tag{3.16}$$

Together (3.15) and (3.16) gives the desired result.

3.137 1. Assume that f is bounded above in a neighborhood of $\mathbf{x}_0$. Then there exists a ball $B(x_0)$ and constant M such that

$$f(\mathbf{x}) \leq M \text{ for every } \mathbf{x} \in B(x_0)$$

Since f is convex

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{x}_0) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{x}_0) \leq \alpha M + (1-\alpha)f(\mathbf{x}_0) \tag{3.17}$$

2. Given $\mathbf{x} \in B(x_0)$ and $\alpha \in [0,1]$ let

$$\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{x}_0 \tag{3.18}$$

Subtracting $f(\mathbf{x}_0)$ from (3.17) gives

$$f(\mathbf{z}) - f(\mathbf{x}_0) \leq \alpha(M - f(\mathbf{x}_0)) \tag{3.19}$$

Rewriting (3.18)

$$\begin{aligned} (1-\alpha)\mathbf{x}_0 &= \mathbf{z} - \alpha\mathbf{x} \\ (1+\alpha)\mathbf{x}_0 &= \mathbf{z} + \alpha(2\mathbf{x}_0 - \mathbf{x}) \\ \mathbf{x}_0 &= \frac{1}{1+\alpha}\mathbf{z} + \frac{\alpha}{1+\alpha}(2\mathbf{x}_0 - \mathbf{x}) \end{aligned}$$

3. Note that

$$(2\mathbf{x}_0 - \mathbf{x}) = \mathbf{x}_0 - (\mathbf{x} - \mathbf{x}_0) \in B(\mathbf{x}_0)$$

so that

$$f(2\mathbf{x}_0 - \mathbf{x}) \leq M$$

and therefore

$$f(\mathbf{x}_0) \leq \frac{1}{1+\alpha}f(\mathbf{z}) + \frac{\alpha}{1+\alpha}f(2\mathbf{x}_0 - \mathbf{x}) \leq \frac{1}{1+\alpha}f(\mathbf{z}) + \frac{\alpha}{1+\alpha}M$$

which implies

$$\begin{aligned} (1+\alpha)f(\mathbf{x}_0) &\leq f(\mathbf{z}) + \alpha M \\ \alpha(f(\mathbf{x}_0) - M) &\leq f(\mathbf{z}) - f(\mathbf{x}_0) \end{aligned}$$

4. Combined with (3.19) we have

$$\alpha(f(\mathbf{x}_0) - M) \leq f(\mathbf{z}) - f(\mathbf{x}_0) \leq \alpha(M - f(\mathbf{x}_0))$$

or

$$|f(\mathbf{z}) - f(\mathbf{x}_0)| \leq \alpha(M - f(\mathbf{x}_0))$$

and therefore $f(\mathbf{z}) \to f(\mathbf{x}_0)$ as $\mathbf{z} \to \mathbf{x}_0$. f is continuous.

3.138 1. Since S is open, there exists a ball $B_r(\mathbf{x}_1) \subseteq S$. Let $t = 1 + \frac{r}{2}$. Then $\mathbf{x}_0 + t(\mathbf{x}_1 - \mathbf{x}_0) \in B_r(x_1) \subseteq S$.

2. Let $s = \frac{t-1}{t} r$. The open ball $B_s(\mathbf{x}_1)$ of radius s centered on $\mathbf{x}_1$ is contained in T. Therefore T is a neighborhood of $\mathbf{x}_1$.

3. Since f is convex, for every $\mathbf{y} \in T$

$$f(\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{z}) \leq (1-\alpha)M + \alpha f(\mathbf{z}) \leq M + f(\mathbf{z})$$

Therefore f is bounded on T.

3.139 The previous exercise showed that f is locally bounded from above for every $\mathbf{x} \in S$. To show that it is also locally bounded from below, choose some $\mathbf{x}_0 \in S$. There exists some $B(\mathbf{x}_0$ and M such that

$$f(\mathbf{x}) \leq M \text{ for every } \mathbf{x} \in B(\mathbf{x}_0)$$

Choose some $x_1 \in B(\mathbf{x}_0)$ and let $\mathbf{x}_2 = 2\mathbf{x}_0 - \mathbf{x}_1$. Then

$$\mathbf{x}_2 = 2\mathbf{x}_0 - \mathbf{x}_1 = \mathbf{x}_0 - (\mathbf{x}_1 - \mathbf{x}_0) \in B(\mathbf{x}_0)$$

and $f(\mathbf{x}_2) \leq M$. Since F is convex

$$f(\mathbf{x}) \leq \frac{1}{2} f(\mathbf{x}_1) + \frac{1}{2} f(\mathbf{x}_2)$$

and therefore

$$f(\mathbf{x}_1) \geq 2f(\mathbf{x}) - f(\mathbf{x}_2)$$

Since $f(\mathbf{x}_2) \leq M$, $-f(\mathbf{x}_2) \geq -M$ and therefore

$$f(\mathbf{x}_1) \geq 2f(\mathbf{x}) - M$$

so that f is bounded from below.

3.140 Let f be a convex function defined on an open convex set S in a normed linear space, which is bounded from above in a neighborhood of a single point $\mathbf{x}_0 \in S$. By Exercise 3.138, f is bounded above at every $\mathbf{x} \in S$. This implies (Exercise 3.137) that f is continuous at every $\mathbf{x} \in S$.

3.141 Without loss of generality, assume $\mathbf{0} \in S$. Assume S has dimension n and let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be a basis for the subspace containing S. Choose some $\lambda > 0$ small enough so that

$$U = \text{conv}\,\{\mathbf{0}, \lambda\mathbf{x}_1, \lambda\mathbf{x}_2, \ldots, \lambda x_n\} \subseteq S$$

Any $\mathbf{x} \in U$ is a convex combination of the points $\mathbf{0}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and so there exists $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n \geq 0$, $\sum \alpha_i = 1$ such that $\mathbf{x} = \alpha_0\mathbf{0} + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$. By Jensen's inequality

$$\begin{aligned} f(\mathbf{x}) = f(\alpha_0\mathbf{0} + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n) &\leq \alpha_0 f(\mathbf{0}) + \alpha_1 f(\mathbf{x}_1) + \cdots + \alpha_n f(\mathbf{x}_n) \\ &\leq \max\{\, f(\mathbf{0}), f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\,\} \end{aligned}$$

Therefore, f is bounded above on a neighbourhood of some $\mathbf{x}_0 \in \text{int}\, U$ (which is nonempty by Exercise 1.229). By Proposition 3.8, f is continuous on S.

3.142 Clearly, if f is convex, it is locally convex at every $\mathbf{x} \in S$, where S is the required neighborhood. To prove the converse, assume to the contrary that f is locally convex at every $\mathbf{x} \in S$ but it is not globally convex. That is, there exists $\mathbf{x}_1, \mathbf{x}_2 \in S$ such that

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) > \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$

Let

$$h(t) = f\big(t\mathbf{x}_1 + (1-t)\mathbf{x}_2\big)$$

Local convexity implies that f is continuous at every $\mathbf{x} \in S$ (Corollary 3.8.1), and therefore continuous on S. Therefore, h is continuous on $[0,1]$. By the continuous maximum theorem (Theorem 2.3),

$$T = \arg \max_{\mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]} h(t)$$

is nonempty and compact. Let $t_0 = \max T$. For every $\epsilon > 0$,

$$h(t_0 - \epsilon) \leq h(t_0) \text{ and } h(t_0 + \epsilon) < h(t_0)$$

Let

$$\mathbf{x}_0 = t_0 \mathbf{x}_1 + (1-t_0)\mathbf{x}_2 \text{ and } \mathbf{x}_\epsilon = (t_0 + \epsilon)\mathbf{x}_1 + (1 - t_0 - \epsilon)\mathbf{x}_2$$

Every neighborhood V of $\mathbf{x}_0$ contains $\mathbf{x}_{-\epsilon}, \mathbf{x}_\epsilon \in [\mathbf{x}_1, \mathbf{x}_2]$ with

$$\frac{1}{2} f(\mathbf{x}_{-\epsilon}) + \frac{1}{2} f(\mathbf{x}_\epsilon) = \frac{1}{2} h(t_0 - \epsilon) + \frac{1}{2} h(t_0 + \epsilon) < h(t_0) = f(\mathbf{x}_0) = f\left(\frac{1}{2}\mathbf{x}_{-\epsilon} + \frac{1}{2}\mathbf{x}_\epsilon\right)$$

contradicting the local convexity of f at $\mathbf{x}_0$.

3.143 Assume f is quasiconcave. That is for every $\mathbf{x}_1, \mathbf{x}_2 \in S$ and $0 \leq \alpha \leq 1$

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \min\{f(\mathbf{x}_1), (\mathbf{x}_2)\}$$

Multiplying through by -1 reverses the inequality so that

$$-f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq -\min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\} = \max\{-f(\mathbf{x}_1), -f(\mathbf{x}_2)\}$$

which shows that $-f$ is quasiconvex. The converse follows analogously.

3.144 Assume f is concave, that is

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \text{ for every } \mathbf{x}_1, \mathbf{x}_2 \in S \text{ and } 0 \leq \alpha \leq 1$$

Without loss of generality assume that $f(\mathbf{x}_1) \leq f(\mathbf{x}_2)$. Then

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_1) = f(\mathbf{x}_1) = \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

f is quasiconcave.

3.145 Let $f \colon \Re \to \Re$. Choose any x_1, x_2 in $\Re$ with $x_1 < x_2$. If f is increasing, then

$$f(x_1) \leq f(\alpha x_1 + (1-\alpha)x_2) \leq f(x_2)$$

for every $0 \leq \alpha \leq 1$. The first inequality implies that

$$f(x_1) = \min\{f(x_1), f(x_2)\} \leq f(\alpha x_1 + (1-\alpha)x_2)$$

so that f is quasiconcave. The second inequality implies that

$$f(\alpha x_1 + (1-\alpha)x_2) \le \max\{f(x_1), f(x_2)\} = f(x_2)$$

so that f is also quasiconvex.

Conversely, if f is decreasing

$$f(x_1) \ge f(\alpha x_1 + (1-\alpha)x_2) \ge f(x_2)$$

for every $0 \le \alpha \le 1$. The first inequality implies that

$$f(x_1) = \max\{f(x_1), f(x_2)\} \ge f(\alpha x_1 + (1-\alpha)x_2)$$

so that f is quasiconvex. The second inequality implies that

$$f(\alpha x_1 + (1-\alpha)x_2) \le \max\{f(x_1), f(x_2)\} = f(x_2)$$

so that f is also quasiconcave.

3.146

$$\precsim_f(c) = \{\, \mathbf{x} \in X : f(\mathbf{x}) \le a \,\} = \{\mathbf{x} \in X : -f(\mathbf{x}) \ge -c\} = \succsim_{-f}(-c)$$

3.147 For given c and m, choose any $\mathbf{p}_1$ and $\mathbf{p}_2$ in $\precsim_v(c)$. For any $0 \le \alpha \le 1$, let $\bar{\mathbf{p}} = \alpha\mathbf{p}_1 + (1-\alpha)\mathbf{p}_2$. The key step is to show that any commodity bundle $\mathbf{x}$ which is affordable at $\bar{\mathbf{p}}$ is also affordable at either $\mathbf{p}_1$ or $\mathbf{p}_2$. Assume that $\mathbf{x}$ is affordable at $\bar{\mathbf{p}}$, that is $\mathbf{x}$ is in the budget set

$$\mathbf{x} \in X(\bar{\mathbf{p}}, m) = \{\, \mathbf{x} : \bar{\mathbf{p}}\mathbf{x} \le m \,\}$$

To show that $\mathbf{x}$ is affordable at either $\mathbf{p}_1$ or $\mathbf{p}_2$, that is

$$\mathbf{x} \in X(\mathbf{p}_1, m) \text{ or } \mathbf{x} \in X(\mathbf{p}_2, m)$$

assume to the contrary that

$$\mathbf{x} \notin X(\mathbf{p}_1, m) \text{ and } \mathbf{x} \notin X(\mathbf{p}_2, m)$$

This implies that

$$\mathbf{p}_1\mathbf{x} > m \text{ and } \mathbf{p}_2\mathbf{x} > m$$

so that

$$\alpha\mathbf{p}_1\mathbf{x} > \alpha m \text{ and } (1-\alpha)\mathbf{p}_2 > (1-\alpha)m$$

Summing these two inequalities

$$\bar{\mathbf{p}}\mathbf{x} = (\alpha\mathbf{p}_1 + (1-\alpha)\mathbf{p}_2)\mathbf{x} > m$$

contradicting the assumption that $\mathbf{x} \in X(\bar{\mathbf{p}}, m)$. We conclude that

$$X(\bar{\mathbf{p}}, m) \subseteq X(\mathbf{p}_1, m) \cup X(\mathbf{p}_2, m)$$

Now

$$\begin{aligned} v(\bar{p}, m) &= \sup\{\, u(\mathbf{x}) : \mathbf{x} \in X(\bar{\mathbf{p}}, m) \,\} \\ &\le \sup\{\, u(\mathbf{x}) : \mathbf{x} \in X(\mathbf{p}_1, m) \cup X(\mathbf{p}_2, m) \,\} \\ &\le c \end{aligned}$$

Therefore $\bar{\mathbf{p}} \in \precsim_v(c)$ for every $0 \le \alpha \le 1$. Thus, $\precsim_v(c)$ is convex and so v is quasiconvex (Exercise 3.146).

3.148 Since f is quasiconcave

$$f(\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2) \geq \min\{f(\mathbf{x}^1), f(\mathbf{x}^2)\} \text{ for every } \mathbf{x}^1, \mathbf{x}^2 \in S \text{ and } 0 \leq \alpha \leq 1$$

Since g is increasing

$$g\big(f(\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2)\big) \geq g\big(\big(\min\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}\big)\big) \geq \min\{g\big(f(\mathbf{x}^1)\big), g\big(f(\mathbf{x}^2)\big)\}$$

$g \circ f$ is quasiconcave.

3.149 When $\rho \geq 1$, the function

$$h(\mathbf{x}) = \alpha_1 x_1^\rho + \alpha_2 x_2^\rho + \dots \alpha_n x_n^\rho$$

is convex (Example 3.58) as is $y^{1/\rho}$. Therefore

$$f(\mathbf{x}) = \big(h(\mathbf{x})\big)^{1/\rho}$$

is an increasing convex function of a convex function and is therefore convex (Exercise 3.133).

3.150 f is a monotonic transformation of the concave function $h(\mathbf{x}) = \mathbf{x}$.

3.151 By Exercise 3.39, there exist linear functionals $\hat{f}$ and $\hat{g}$ and scalars b and c such that

$$f(\mathbf{x}) = \hat{f}(\mathbf{x}) + b \text{ and } g(\mathbf{x}) = \hat{g}(\mathbf{x}) + c$$

The upper contour set

$$\begin{aligned}
\succsim_h(a) &= \{\, x \in S : h(\mathbf{x}) \geq a \,\} \\
&= \{\, x \in S : \frac{\hat{f}(x) + b}{\hat{g}(x) + c} \geq a \,\} \\
&= \{\, x \in \Re^n_+ : \hat{f}(\mathbf{x}) + b \geq a\hat{g}(\mathbf{x}) + ac \,\} \\
&= \{\, x \in \Re^n_+ : \hat{f}(\mathbf{x}) - a\hat{g}(\mathbf{x}) \geq b - ac \,\}
\end{aligned}$$

which is a halfspace in X and therefore convex. Similarly, the lower contour set

$$\precsim_h(a) = \{\, x \in S : h(\mathbf{x}) \geq a \,\}$$

is also a halfspace and hence convex. Therefore h is both quasiconcave and quasiconvex.

3.152 For $a \leq 0$

$$\succsim(a) = \{\, x \in S : h(\mathbf{x}) \geq 0 \,\} = S$$

which is convex. For $a > 0$

$$\begin{aligned}
\succsim_h(a) &= \{\, x \in S : h(\mathbf{x}) \geq a \,\} \\
&= \{\, x \in S : \frac{f(\mathbf{x})}{g(\mathbf{x})} \geq a \,\} \\
&= \{\, x \in S : f(\mathbf{x}) \geq ag(\mathbf{x}) \,\} \\
&= \{\, x \in S : f(\mathbf{x}) - ag(\mathbf{x}) \geq 0 \,\}
\end{aligned}$$

is convex since $f - ag = f + a(-g)$ is concave (Exercises 3.124 and 3.131). Since $\succsim_h(a)$ is convex for every a, h is quasiconcave.

3.153

$$h(\mathbf{x}) = \frac{f(\mathbf{x})}{\hat{g}(\mathbf{x})}$$

where $\hat{g} = 1/g$ is positive and convex by Exercise 3.135. By the previous exercise, h is quasiconcave.

3.154 Let $F = \log f$. If F is concave, $f(\mathbf{x}) = e^{F(\mathbf{x})}$ is an increasing function of (quasi)concave function, and hence is quasiconcave (Exercise 3.148).

3.155 Let

$$F(\mathbf{x}) = \log f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i \log f_i(\mathbf{x})$$

As the sum of concave functions, F is concave (Exercise 3.131). By the previous exercise, f is quasiconcave.

3.156 Assume $\mathbf{x}_1$, $\mathbf{x}_2$ and $\bar{\mathbf{x}}$ are optimal solutions for $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\bar{\boldsymbol{\theta}} = \alpha\boldsymbol{\theta}_1 + (1-\alpha)\boldsymbol{\theta}_2$ respectively. That is

$$\begin{aligned} f(\mathbf{x}_1, \boldsymbol{\theta}_1) &= v(\boldsymbol{\theta}_1) \\ f(\mathbf{x}_2, \boldsymbol{\theta}_2) &= v(\boldsymbol{\theta}_2) \\ f(\bar{\mathbf{x}}, \bar{\boldsymbol{\theta}}) &= v(\bar{\boldsymbol{\theta}}) \end{aligned}$$

Since f is convex in $\boldsymbol{\theta}$

$$\begin{aligned} v(\bar{\boldsymbol{\theta}}) &= f(\bar{\mathbf{x}}, \bar{\boldsymbol{\theta}}) \\ &= f(\bar{\mathbf{x}}, \alpha\boldsymbol{\theta}_1 + (1-\alpha)\boldsymbol{\theta}_2) \\ &\leq \alpha f(\bar{\mathbf{x}}, \boldsymbol{\theta}_1) + (1-\alpha) f(\mathbf{x}^*, \boldsymbol{\theta}_2) \\ &\leq \alpha f(\mathbf{x}_1, \boldsymbol{\theta}_1) + (1-\alpha) f(\mathbf{x}_2, \boldsymbol{\theta}_2) \\ &= \alpha v(\boldsymbol{\theta}_1) + (1-\alpha) v(\boldsymbol{\theta}_2) \end{aligned}$$

v is convex.

3.157 Assume to the contrary that $\mathbf{x}_1$ and $\mathbf{x}_2$ are distinct optimal solutions, that is $\mathbf{x}_1, \mathbf{x}_2 \in \varphi(\boldsymbol{\theta})$, $\mathbf{x}^1 \neq \mathbf{x}^2$, for some $\boldsymbol{\theta} \in \Theta^*$, so that

$$f(\mathbf{x}_1, \boldsymbol{\theta}) = f(\mathbf{x}_2, \boldsymbol{\theta}) = v(\boldsymbol{\theta}) \geq f(\mathbf{x}, \boldsymbol{\theta}) \quad \text{for every} \quad \mathbf{x} \in G(\boldsymbol{\theta})$$

Let $\bar{\mathbf{x}} = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ for $\alpha \in (0,1)$. Since $G(\boldsymbol{\theta})$ is convex, $\bar{\mathbf{x}}$ is feasible. Since f is strictly quasiconcave

$$f(\bar{\mathbf{x}}, \boldsymbol{\theta}) > \min\{\, f(\mathbf{x}_1, \boldsymbol{\theta}), f(\mathbf{x}_2, \boldsymbol{\theta}) \,\} = v(\boldsymbol{\theta})$$

contradicting the optimality of $\mathbf{x}_1$ and $\mathbf{x}_2$. We conclude that $\varphi(\boldsymbol{\theta})$ is single-valued for every $\boldsymbol{\theta} \in \Theta^*$. In other words, φ is a function.

3.158 1. The value function is

$$v(x_0) = \sup_{\mathbf{x} \in \Gamma(x_0)} U(\mathbf{x})$$

where

$$U(\mathbf{x}) = \sum_{t=0}^{\infty} \beta^t f(x_t, x_{t+1})$$

and

$$\Gamma(x_0) = \{\mathbf{x} \in X^\infty : x_{t+1} \in G(x_t), t = 0, 1, 2, \dots\}$$

Since an optimal policy exists (Exercise 2.125), the maximum is attained and

$$v(x_0) = \max_{\mathbf{x} \in \Gamma(x_0)} U(\mathbf{x}) \tag{3.20}$$

It is straightforward to show that

- $U(\mathbf{x})$ is strictly concave and
- $\Gamma(x_0)$ is convex

Applying the Concave Maximum Theorem (Theorem 3.1) to (3.20), we conclude that the value function v is strictly concave.

2. Assume to the contrary that $\mathbf{x}'$ and $\mathbf{x}''$ are distinct optimal plans, so that

$$v(x_0) = U(\mathbf{x}') = U(\mathbf{x}'')$$

Let $\bar{\mathbf{x}} = \alpha\mathbf{x}' + (1-\alpha)\mathbf{x}''$. Since $\Gamma(x_0)$ is convex, $\bar{\mathbf{x}}$ is feasible and

$$U(\bar{\mathbf{x}}) > \alpha U(\mathbf{x}') + (1-\alpha)U(\mathbf{x}'') = U(\mathbf{x}')$$

which contradicts the optimality of $\mathbf{x}'$. We conclude that the optimal plan is unique.

3.159 We observe that

- $u(F(k) - y)$ is supermodular in y (Exercise 2.51)
- $u(F(k) - y)$ displays strictly increasing differences in (k, y) (Exercise 3.129)
- $G(k) = [0, F(k)]$ is increasing.

Applying Exercise 2.126, we can conclude that the optimal policy $(k_0, k_1^*, k_2^*, \dots)$ is a monotone sequence. Since X is compact, $\mathbf{k}^*$ is a bounded monotone sequence, which converges monotonically to some steady state k^* (Exercise 1.101).

3.160 Suppose there exists $(\mathbf{x}^*, \mathbf{y}^*) \in X \times Y$ such that

$$f(\mathbf{x}, \mathbf{y}^*) \le f(\mathbf{x}^*, \mathbf{y}^*) \le f(\mathbf{x}^*, \mathbf{y}) \text{ for every } \mathbf{x} \in X \text{ and } \mathbf{y} \in Y$$

Let $v = f(\mathbf{x}^*, \mathbf{y}^*)$. Since

$$f(\mathbf{x}, \mathbf{y}^*) \le v \text{ for every } \mathbf{x} \in X$$
$$\max_{\mathbf{x} \in X} f(x, \mathbf{y}^*) \le v$$

and therefore

$$\min_{\mathbf{y} \in Y} \max_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}) \le \max_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}^*) \le v$$

Similarly

$$\max_{\mathbf{x} \in X} \min_{\mathbf{y} \in y} f(\mathbf{x}, \mathbf{y}) \ge v$$

Combining the last two inequalities, we have

$$\max_{\mathbf{x} \in X} \min_{\mathbf{y} \in y} f(\mathbf{x}, \mathbf{y}) \ge v \ge \min_{\mathbf{y} \in Y} \max_{\mathbf{x} \in X} f(x, \mathbf{y})$$

Together with (3.28), this implies equality

$$\max_{\mathbf{x}\in X}\min_{\mathbf{y}\in Y} f(\mathbf{x},\mathbf{y}) = \min_{\mathbf{y}\in Y}\max_{\mathbf{x}\in X} f(\mathbf{x},\mathbf{y})$$

Conversely, suppose that

$$\max_{\mathbf{x}\in X}\min_{\mathbf{y}\in Y} f(\mathbf{x},\mathbf{y}) = v = \min_{\mathbf{y}\in Y}\max_{\mathbf{x}\in X} f(\mathbf{x},\mathbf{y})$$

The function

$$g(\mathbf{x}) = \min_{\mathbf{y}\in Y} f(\mathbf{x},\mathbf{y})$$

is a continuous function (Theorem 2.3) on a compact set X. By the Weierstrass theorem (Theorem 2.2), there exists $\mathbf{x}^*$ which maximizes g on X, that is

$$g(\mathbf{x}^*) = \min_{\mathbf{y}\in Y} f(\mathbf{x}^*,\mathbf{y}) = \max_{\mathbf{x}\in X} g(\mathbf{x}) = \max_{\mathbf{x}\in X}\min_{\mathbf{y}\in Y} f(\mathbf{x},\mathbf{y}) = v$$

which implies that

$$f(\mathbf{x}^*,\mathbf{y}) \geq v \text{ for every } \mathbf{y}\in Y$$

Similarly, there exists $\mathbf{y}\in Y$ such that

$$f(\mathbf{x},\mathbf{y}^*) \leq v \text{ for every } \mathbf{x}\in X$$

Combining these inequalities, we have

$$f(\mathbf{x},\mathbf{y}^*) \leq v \leq f(\mathbf{x}^*,\mathbf{y}) \text{ for every } \mathbf{x}\in X \text{ and } \mathbf{y}\in Y$$

In particular, we have

$$f(\mathbf{x}^*,\mathbf{y}^*) \leq v \leq f(\mathbf{x}^*,\mathbf{y}^*)$$

so that $v = f(\mathbf{x}^*,\mathbf{y}^*)$ as required.

3.161 For any $x\in X$ and $y\in Y$, let

$$g(x) = \min_{y\in Y} f(x,y) \text{ and } h(y) = \max_{x\in X} f(x,y)$$

Then

$$g(x) = \min_{y\in Y} f(x,y) \leq \max_{x\in X} f(x,y) = h(y)$$

and therefore

$$\max_{x\in X} g(x) \leq \max_{y\in Y} h(y)$$

That is

$$\max_x \min_y f(x,y) \leq f_y \max_x f(x,y)$$

3.162 Clearly $f(x) = x^a$ his homogeneous of degree a. Conversely assume f is homogeneous of degree a, that is

$$f(tx) = t^a f(x)$$

Letting $x = 1$

$$f(t) = t^a f(1)$$

Setting $f(1) = A \in \Re$ and interchanging x and t yields the result.

3.163

$$\begin{aligned} f(t\mathbf{x}) &= (a_1(tx_1)^\rho + a_2(tx_2)^\rho + \dots a_n(tx_n)^\rho)^{1/\rho} \\ &= t\,(a_1x_1^\rho + a_2x_2^\rho + \dots a_nx_n^\rho)^{1/\rho} \\ &= tf(\mathbf{x}) \end{aligned}$$

3.164 For $\beta \in \Re_{++}$

$$h(\beta t) = f(\beta t\mathbf{x}_0) = \beta^k f(t\mathbf{x}_0) = \beta^k h(t)$$

3.165 Suppose that $\mathbf{x}^*$ minimizes the cost of producing output y at prices $\mathbf{w}$. That is

$$\mathbf{w}^T\mathbf{x}^* \leq \mathbf{w}^T\mathbf{x} \qquad \text{for every } \ \mathbf{x} \in V(y)$$

It follows that

$$t\mathbf{w}^T\mathbf{x}^* \leq t\mathbf{w}^T\mathbf{x} \qquad \text{for every } \ \mathbf{x} \in V(y)$$

for every $t > 0$, verifying that $\mathbf{x}^*$ minimizes the cost of producing y at prices $t\mathbf{w}$. Therefore

$$c(t\mathbf{w}, y) = (t\mathbf{w})\mathbf{x}^* = t(\mathbf{w}^T\mathbf{x}^*) = tc(\mathbf{w}, y)$$

$c(\mathbf{w}, y)$ homogeneous of degree one in input prices $\mathbf{w}$.

3.166 For given prices $\mathbf{w}$, let $\mathbf{x}^*$ minimize the cost of producing one unit of output, so that $c(\mathbf{w}, 1) = \mathbf{w}^T\mathbf{x}^*$. Clearly $f(\mathbf{x}^*) = 1$ where f is the production function.

Now consider any output y. Since f is homogeneous

$$f(y\mathbf{x}^*) = yf(\mathbf{x}^*) = y$$

Therefore $y\mathbf{x}^*$ is sufficient to produce y, so that

$$c(\mathbf{w}, y) \leq \mathbf{w}^T(y\mathbf{x}^*) = y\mathbf{w}^T\mathbf{x}^* = yc(\mathbf{w}, 1)$$

Suppose that

$$c(\mathbf{w}, y) < \mathbf{w}^T(y\mathbf{x}^*) = yc(\mathbf{w}, 1)$$

Then there exists $\mathbf{x}'$ such that $f(\mathbf{x}') = y$ and

$$\mathbf{w}^T\mathbf{x}' < \mathbf{w}^T(y\mathbf{x}^*)$$

which implies that

$$\mathbf{w}^T\left(\frac{\mathbf{x}'}{y}\right) < \mathbf{w}^T\mathbf{x}^* = c(\mathbf{w}, 1)$$

Since f is homogeneous

$$f\left(\frac{\mathbf{x}'}{y}\right) = \frac{1}{y}f(\mathbf{x}') = 1$$

Therefore, $\mathbf{x}'$ is a lower cost method of producing one unit of output, contradicting the definition of $\mathbf{x}^*$. We conclude that

$$c(\mathbf{w}, y) = yc(\mathbf{w}, 1)$$

$c(\mathbf{w}, y)$ is homogeneous of degree one in y.

3.167 If the consumer's demand is invariant to proportionate changes in all prices and income, so also will the derived utility. More formally, suppose that $\mathbf{x}^*$ maximizes utility at prices $\mathbf{p}$ and income m, that is

$$\mathbf{x}^* \succsim \mathbf{x} \qquad \text{for every } \ \mathbf{x} \in X(\mathbf{p}, m)$$

Then

$$v(\mathbf{p}, m) = u(\mathbf{x}^*)$$

Since $X(t\mathbf{p}, tm) = X(\mathbf{p}, m)$

$$\mathbf{x}^* \succsim \mathbf{x} \qquad \text{for every } \ \mathbf{x} \in X(t\mathbf{p}, tm)$$

and

$$v(t\mathbf{p}, tm) = u(\mathbf{x}^*) = v(\mathbf{p}, m)$$

3.168 Assume f is homogeneous of degree one, so that

$$f(t\mathbf{x}) = tf(\mathbf{x}) \qquad \text{for every } \ t > 0$$

Let $(\mathbf{x}, y) \in \text{epi } f$, so that

$$f(\mathbf{x}) \leq y$$

For any $t > 0$

$$f(t\mathbf{x}) = tf(\mathbf{x}) \leq ty$$

which implies that $(t\mathbf{x}, ty) \in \text{epi } f$. Therefore epi f is a cone.

Conversely assume epi f is a cone. Let $\mathbf{x} \in S$ and define $y = f(\mathbf{x})$. Then $(\mathbf{x}, y) \in \text{epi } f$ and therefore $(t\mathbf{x}, ty) \in \text{epi } f$ so

$$f(t\mathbf{x}) \leq ty$$

Now suppose to the contrary that

$$f(t\mathbf{x}) = z < ty = tf(\mathbf{x}) \tag{3.21}$$

Then $(t\mathbf{x}, z) \in \text{epi } f$. Since epi f is a cone, we must have $(\mathbf{x}, z/t) \in \text{epi } f$ so that

$$f(\mathbf{x}) \leq \frac{z}{t}$$

and

$$tf(\mathbf{x}) \leq z = f(t\mathbf{x})$$

contradicting (3.21). We conclude that

$$f(t\mathbf{x}) = tf(\mathbf{x}) \text{ for every } \ t > 0$$

3.169 Take any $\mathbf{x}_1$ and $\mathbf{x}_2$ in S and let

$$y_1 = f(\mathbf{x}_1) > 0 \text{ and } \; y_2 = f(\mathbf{x}_2) > 0$$

Since f is homogeneous of degree one,

$$f\left(\frac{\mathbf{x}_1}{y_1}\right) = f\left(\frac{\mathbf{x}_2}{y_2}\right) = 1$$

Since f is quasiconcave

$$f\left(\alpha\frac{\mathbf{x}_1}{y_1} + (1-\alpha)\frac{\mathbf{x}_2}{y_2}\right) \geq 1$$

for every $0 \leq \alpha \leq 1$. Choose $\alpha = y_1/(y_1+y_2)$ so that $(1-\alpha) = y_2/(y_1+y_2)$. Then

$$f\left(\frac{\mathbf{x}_1}{y_1+y_2} + \frac{\mathbf{x}_2}{y_1+y_2}\right) \geq 1$$

Again using the homogeneity of f, this implies

$$f\left(\mathbf{x}_1 + \mathbf{x}_2\right) \geq y_1 + y_2 = f(\mathbf{x}_1) + f(\mathbf{x}_2)$$

3.170 Let $f \in F(S)$ be a strictly positive definite, quasiconcave functional which is homogeneous of degree one. For any $\mathbf{x}_1$, $\mathbf{x}_2$ in S and $0 \leq \alpha \leq 1$ $\alpha\mathbf{x}_1$, $(1-\alpha)\mathbf{x}_2$ in S and therefore

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq f(\alpha\mathbf{x}_1) + f((1-\alpha)\mathbf{x}_2)$$

since f is superadditive (Exercise 3.169). But

$$f(\alpha\mathbf{x}_1) = \alpha f(\mathbf{x}_1)$$
$$f((1-\alpha)\mathbf{x}_2) = (1-\alpha)f(\mathbf{x}_2)$$

by homogeneity. Substituting in (3.21), we conclude that

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha)f((1-\alpha)\mathbf{x}_2)$$

f is concave.

3.171 Assume that f is strictly positive definite, quasiconcave and homogeneous of degree k, $0 < k < 1$. Define

$$h(\mathbf{x}) = \left(f(\mathbf{x})\right)^{1/k}$$

Then h is quasiconcave (Exercise 3.148. Further, for every $t > 0$

$$\begin{aligned} h(t\mathbf{x}) &= \left(f(t\mathbf{x})\right)^{1/k} \\ &= \left(t^k f(\mathbf{x})\right)^{1/k} \\ &= t\left(f(\mathbf{x})\right)^{1/k} \\ &= th(\mathbf{x}) \end{aligned}$$

so that h is homogeneous of degree 1. By Exercise 3.170, h is concave.

$$f(\mathbf{x}) = \left(h(\mathbf{x})\right)^k$$

That is $f = g \circ h$ where

$$g(y) = y^k$$

is monotone and concave provided $k \leq 1$. By Exercise 3.133, $f = g \circ h$ is concave.

3.172 Continuity is a necessary and sufficient condition for the existence of a utility function representing $\succsim$ (Remark 2.9).

Suppose u represents the homothetic preference relation $\succsim$. For any $\mathbf{x}_1, \mathbf{x}_2 \in S$

$$u(\mathbf{x}_1) = u(\mathbf{x}_2) \implies \mathbf{x}_1 \sim \mathbf{x}_2 \implies t\mathbf{x}_1 \sim t\mathbf{x}_2 \implies u(t\mathbf{x}_1) = u(t\mathbf{x}_2) \text{ for every } \ t > 0$$

Conversely, if u is a homothetic functional,

$$\mathbf{x}_1 \sim \mathbf{x}_2 \implies u(\mathbf{x}_1) = u(\mathbf{x}_2) \implies u(t\mathbf{x}_1) = u(t\mathbf{x}_2) \implies t\mathbf{x}_1 \sim t\mathbf{x}_2 \text{ for every } \ t > 0$$

3.173 Suppose that $f = g \circ h$ where g is strictly increasing and h is homogeneous of degree k. Then

$$\hat{h}(\mathbf{x}) = \big(h(\mathbf{x})\big)^{1/k}$$

is homogeneous of degree one and $f = \hat{g} \circ \hat{h}$ where

$$\hat{g}(y) = g\left(y^k\right))$$

is increasing.

3.174 Assume $\mathbf{x}^1, \mathbf{x}^2 \in S$ with

$$f(\mathbf{x}^1) = g(h(\mathbf{x}^1)) = g(h\mathbf{x}^2)) = f(\mathbf{x}^2)$$

Since g is strictly increasing, this implies that

$$h(\mathbf{x}^1) = h(\mathbf{x}^2)$$

Since h is homogeneous

$$h(t\mathbf{x}^1) = t^k h(\mathbf{x}^1) = t^k h(\mathbf{x}^2) = h(t\mathbf{x}^2)$$

for some k. Therefore

$$f(t\mathbf{x}^1) = g(h(t\mathbf{x}^1)) = g(h(t\mathbf{x}^2)) = f(t\mathbf{x}^2)$$

3.175 Let $\mathbf{x}_0 \neq \mathbf{0}$ be any point in S, and define $g \colon \Re \to \Re$ by

$$g(\alpha) = f(\alpha \mathbf{x}_0)$$

Since f is strictly increasing, so is g and therefore g has a strictly increasing inverse g^{-1}. Let $h = g^{-1} \circ f$ so that $f = g \circ h$.

We need to show that h is homogeneous. For any $\mathbf{x} \in S$, there exists α such that

$$g(\alpha) = f(\alpha \mathbf{x}_0) = f(\mathbf{x})$$

that is $\alpha = h(\mathbf{x}) = g^{-1}(f(\mathbf{x}))$. Since f is homothetic

$$g(t\alpha) = f(t\alpha\mathbf{x}_0)f(t\mathbf{x}) \text{ for every } \ t > 0$$

and therefore

$$h(t\mathbf{x}) = g^{-1}(f(t\mathbf{x})) = g^{-1}(f(t\alpha\mathbf{x}_0)) = g^{-1}g(t\alpha) = t\alpha = th(\mathbf{x})$$

h is homogeneous of degree one.

3.176 Let f be the production function. If f is homothetic, there exists (Exercise 3.175) a linearly homogeneous function h and strictly increasing function g such that $f = g \circ h$.

$$\begin{aligned} c(\mathbf{w}, y) &= \min_{\mathbf{x}} \{ \mathbf{w}^T \mathbf{x} : f(\mathbf{x}) \geq y \} \\ &= \min_{\mathbf{x}} \{ \mathbf{w}^T \mathbf{x} : g(h(\mathbf{x})) \geq y \} \\ &= \min_{\mathbf{x}} \{ \mathbf{w}^T \mathbf{x} : h(\mathbf{x}) \geq g^{-1}(y) \} \\ &= g^{-1}(y) c(\mathbf{w}, 1) \end{aligned}$$

by Exercise 3.166.

3.177 Let $f \colon S \to \Re$ be positive, strictly increasing, homothetic and quasiconcave. By Exercise 3.175, there exists a linearly homogeneous function $h : S \to \Re$ and strictly increasing function $g \in F(R)$ such that $f = g \circ h$. $h = g^{-1} \circ f$ is positive, quasiconcave (Exercise 3.148) and homogeneous of degree one. By Proposition 3.12, h is concave and therefore $f = g \circ h$ is concavifiable.

3.178 Since $H_f(c)$ is a supporting hyperplane to S at $\mathbf{x}_0$, then

$$f(\mathbf{x}_0) = c$$

and either

$$f(\mathbf{x}) \geq c = f(\mathbf{x}_0) \text{ for every } \mathbf{x} \in S$$

or

$$f(\mathbf{x}) \leq c = f(\mathbf{x}_0) \text{ for every } \mathbf{x} \in S$$

3.179 Suppose to the contrary that $\mathbf{y} = (h, q) \in \text{int } A \cap B$. Then $\mathbf{y} \succsim \mathbf{y}^*$. By strict convexity

$$\mathbf{y}^\alpha = \alpha \mathbf{y} + (1 - \alpha) \mathbf{y}^* \succ \mathbf{y}^* \text{ for every } \alpha \in (0, 1)$$

Since $\mathbf{y} \in \text{int } A$, $\mathbf{y}^\alpha \in A$ for α sufficiently small. That is, there exists some α such that $\mathbf{y}^\alpha$ is feasible and $\mathbf{y}^\alpha \succ \mathbf{y}^*$, contradicting the optimality of $\mathbf{y}^*$.

3.180 For notational simplicity, let f be the linear functional which separates A and B in Example 3.77. $f(\mathbf{y})$ measure the cost of the plan $\mathbf{y} = (h, q)$, that is $f(\mathbf{y}) = wh + pq$.

Assume to the contrary there exists a preferred lifestyle in X, that is there exists some $\mathbf{y} = (h, q) \in X$ such that $\mathbf{y} \succ \mathbf{y}^* = (h^*, q^*)$. Since $\mathbf{y} \in B$, $f(\mathbf{y}) \geq f(\mathbf{y}^*)$ by (3.29). On the other hand, $\mathbf{y} \in X$ which implies that $f(\mathbf{y}) \leq f(\mathbf{y}^*)$. Consequently, $f(\mathbf{y}) = f(\mathbf{y}^*)$.

By continuity, there exists some $\alpha < 1$ such that $\alpha \mathbf{y} \succ \mathbf{y}^*$ which implies that $\alpha \mathbf{y} \in B$. By linearity

$$f(\alpha \mathbf{y}) = \alpha f(\mathbf{y}) < f(\mathbf{y}) = f(\mathbf{y}^*) = \alpha$$

contrary to (3.29). This contradiction establishes that $\mathbf{y}^*$ is the best choice in budget set X.

3.181 By Proposition 3.7, epi f is a convex set in $X \times \Re$ with $(\mathbf{x}_0, f(\mathbf{x}_0))$ a point on its boundary. By Corollary 3.2.2 of the Separating Hyperplane Theorem, there exists linear a functional $\varphi \in (X \times \Re)'$ such that

$$\varphi(\mathbf{x}, y) \geq \varphi(\mathbf{x}_0, f(\mathbf{x}_0)) \text{ for every } (\mathbf{x}, y) \in \text{epi } f \tag{3.22}$$

φ can be decomposed into two components (Exercise 3.47)

$$\varphi(\mathbf{x}, y) = -g(\mathbf{x}) + \alpha y$$

The assumption that $\mathbf{x}_0 \in \text{int } S$ ensures that $\alpha > 0$ and we can normalize so that $\alpha = 1$. Substituting in (3.22)

$$\begin{aligned} -g(\mathbf{x}) + f(\mathbf{x}) &\geq -g(\mathbf{x}_0) + f(\mathbf{x}_0) \\ f(\mathbf{x}) &\geq f(\mathbf{x}_0) + g(\mathbf{x} - \mathbf{x}_0) \end{aligned}$$

for every $\mathbf{x} \in S$.

3.182 By Exercise 3.72, there exists a unique point $\mathbf{x}_0 \in S$ such that

$$(\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x} - \mathbf{x}_0) \geq 0 \text{ for every } \mathbf{x} \in S$$

Define the linear functional (Exercise 3.64)

$$f(\mathbf{x}) = (\mathbf{x}_0 - \mathbf{y})^T \mathbf{x}$$

and let $c = f(\mathbf{x}_0)$. For all $\mathbf{x} \in S$

$$f(\mathbf{x}) - f(\mathbf{x}_0) = f(\mathbf{x} - \mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x} - \mathbf{x}_0) \geq 0$$

and therefore

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) = c \text{ for every } \mathbf{x} \in S$$

Furthermore

$$f(\mathbf{x}_0) - f(\mathbf{y}) = f(\mathbf{x}_0 - \mathbf{y}) = (\mathbf{x}_0 - \mathbf{y})^T(\mathbf{x}_0 - \mathbf{y}) = \|\mathbf{x}_0 - \mathbf{y}\|^2 > 0$$

since $\mathbf{y} \neq \mathbf{x}_0$. Therefore $f(\mathbf{x}_0) > f(\mathbf{y})$ and

$$f(\mathbf{y}) < c \leq f(\mathbf{x}) \text{ for every } \mathbf{x} \in S$$

3.183 If $\mathbf{y} \in \text{b}(S)$, $\mathbf{y} \in \overline{S^c}$ and there exists a sequence of points $\{\mathbf{y}^n\} \in S^c$ converging to $\mathbf{y}$ (Exercise 1.105). That is, there exists a sequence of nonboundary points $\{\mathbf{y}^n\} \notin \overline{S}$ converging to $\mathbf{y}$. For every point $\mathbf{y}^n$, there is a linear functional $g^n \in X^*$ and c^n such that

$$g^n(\mathbf{y}^n) < c^n \leq g^n(\mathbf{x}) \qquad \text{for every } \mathbf{x} \in \overline{S}$$

Define $f^n = g^n / \|g^n\|$. By construction, the sequence of linear functionals f^n belong to the unit ball in X^* (since $\|f\| = 1$). Since X^* is finite dimensional, the unit ball is compact as so f^n has a convergent subsequence with limit f such that

$$f(\mathbf{y}) \leq f(\mathbf{x}) \qquad \text{for every } x \in \overline{S}$$

A fortiori

$$f(\mathbf{y}) \leq f(\mathbf{x}) \qquad \text{for every } x \in S$$

3.184 There are two possible cases.

$\mathbf{y} \notin \overline{S}$ By Exercise 3.182, there exists a hyperplane which separates $\mathbf{y}$ and $\overline{S}$ which *a fortiori* separates $\mathbf{y}$ and S, that is

$$f(\mathbf{y}) \leq f(\mathbf{x}) \qquad \text{for every } \mathbf{x} \in S$$

$\mathbf{y} \in \overline{S}$ Since $\mathbf{y} \notin S$, $\mathbf{y}$ must be a boundary point of S. By the previous exercise, there exists a supporting hyperplane at $\mathbf{y}$, that is there exists a continuous linear functional $f \in X^*$ such that

$$f(\mathbf{y}) \leq f(\mathbf{x}) \qquad \text{for every } \ \mathbf{x} \in S$$

3.185 1. $f(S) \subseteq \Re$.

2. $f(S)$ is convex and hence an interval (Exercise 1.160.
3. $f(S)$ is open in $\Re$ (Proposition 3.2).

3.186 S is nonempty and convex and $\mathbf{0} \notin S$. (Otherwise, there exists $\mathbf{x} \in A$ and $\mathbf{y} \in B$ such that $\mathbf{0} = \mathbf{y} + (-\mathbf{x})$ which implies that $\mathbf{x} = \mathbf{y}$ contradicting the assumption that $A \cap B = \emptyset$.) Thus there exists a continuous linear functional $f \in X^*$ such that

$$f(\mathbf{y} - \mathbf{x}) \geq f(\mathbf{0}) = 0 \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

so that

$$f(\mathbf{x}) \leq f(\mathbf{y}) \text{ for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

Let $c = \sup_{\mathbf{x} \in A} f(\mathbf{x})$. Then

$$f(\mathbf{x}) \leq c \leq f(\mathbf{y}) \text{ for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

By Exercise 3.185, $f(\text{int } A)$ is an open interval in $(-\infty, c]$, hence $f(\text{int } A) \subseteq (-\infty, c)$, so that $f(\mathbf{x}) < c$ for every $\mathbf{x} \in \text{int } A$. Similarly, $f(\text{int } B) > c$ and

$$f(\mathbf{x}) < c < f(\mathbf{y}) \text{ for every } \ \mathbf{x} \in \text{int } A, \mathbf{y} \in \text{int } B$$

3.187 Since $\text{int } A \cap B = \emptyset$, int A and B can be separated. That is, there exists a continuous linear functional $f \in X^*$ and a number c such that

$$f(\mathbf{x}) \leq c \leq f(\mathbf{y}) \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{y} \in \text{int } B$$

which implies that

$$f(\mathbf{x}) \leq c \leq f(\mathbf{y}) \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

since

$$c \leq \inf_{\mathbf{y} \in \text{int } B} f(\mathbf{y}) = \inf_{\mathbf{y} \in B} f(\mathbf{y})$$

Conversely, suppose that A and B can be separated. That is, there exists $f \in X^*$ such that

$$f(\mathbf{x}) \leq c \leq f(\mathbf{y}) \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

Then $f(\text{int } A)$ is an open interval in $[c, \infty)$, which is disjoint from the interval $f(B) \subseteq (-\infty, c]$. This implies that $\text{int } A \cap B = \emptyset$.

3.188 Since $\mathbf{x}_0 \in \text{b}(S)$, $\{\mathbf{x}_0\} \cap \text{int } S = \emptyset$ and $\text{int } S \neq \emptyset$. By Corollary 3.2.1, $\{\mathbf{x}_0\}$ and S can be separated, that is there exist $f \in X^*$ such that

$$f(\mathbf{x}_0) \leq f(\mathbf{x}) \ \text{ for every } \ \mathbf{x} \in S$$

3.189 Let $\mathbf{x} \in C$. Since C is a cone, $\lambda\mathbf{x} \in C$ for every $\lambda \geq 0$ and therefore

$$f(\lambda\mathbf{x}) \geq c$$

or

$$f(\mathbf{x}) \geq c/\lambda \qquad \text{for every} \quad \lambda \geq 0$$

Taking the limit as $\lambda \to \infty$ implies that

$$f(\mathbf{x}) \geq 0 \qquad \text{for every} \quad \mathbf{x} \in C$$

3.190 First note that $\mathbf{0} \in Z$ and therefore $f(\mathbf{0}) = 0 \leq c$ so that $c \geq 0$. Suppose that there exists some $\mathbf{z} \in Z$ for which $f(\mathbf{z}) = \epsilon \neq 0$. By linearity, this implies

$$f(\frac{2c}{\epsilon}\mathbf{z}) = \frac{2c}{\epsilon}f(\mathbf{z}) = 2c > c$$

which contradicts the requirement

$$f(\mathbf{z}) \leq c \text{ for every} \quad \mathbf{z} \in Z$$

3.191 By Corollary 3.2.1, there exists $f \in X^*$ such that

$$f(\mathbf{z}) \leq c \leq f(\mathbf{x}) \qquad \text{for every} \quad \mathbf{x} \in S, \mathbf{z} \in Z$$

By Exercise 3.190

$$f(\mathbf{z}) = 0 \qquad \text{for every} \quad \mathbf{z} \in Z$$

and therefore

$$f(\mathbf{x}) \geq 0 \qquad \text{for every} \quad \mathbf{x} \in S$$

Therefore Z is contained in the hyperplane $H_f(0)$ which separates S from Z.

3.192 Combining Theorem 3.2 and Corollary 3.2.1, there exists a hyperplane $H_f(c)$ such that

$$f(\mathbf{x}) \leq c \leq f(\mathbf{y}) \qquad \text{for every} \quad \mathbf{x} \in A, \mathbf{y} \in B$$

and such that

$$f(\mathbf{x}) < c \leq f(\mathbf{y}) \qquad \text{for every} \quad \mathbf{x} \in \text{int } A, \mathbf{y} \in B$$

Since int $A \neq \emptyset$, there exists some $\mathbf{x} \in \text{int } A$ with $f(\mathbf{x}) < c$. Hence $A \not\subseteq f^{-1}(c) = H_f(c)$.

3.193 Follows directly from the basic separation theorem, since $A = \text{int } A$ and $B = \text{int } B$.

3.194 Let $S = B - A$. Then

1. S is a nonempty, closed, convex set (Exercise 1.203).
2. $\mathbf{0} \notin S$.

There exists a continuous linear functional $f \in X^*$ such that

$$f(\mathbf{x}) \geq c > f(\mathbf{0}) = 0$$

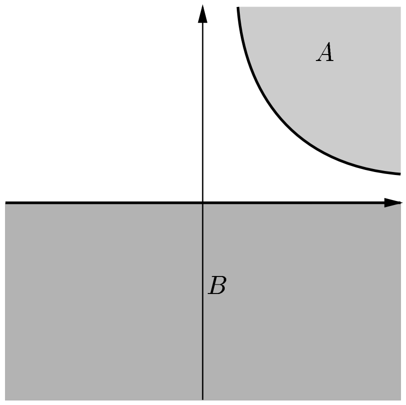

Figure 3.2: A and B cannot be strongly separated.

for every $\mathbf{z} \in S$ (Exercise 3.182). For every $\mathbf{x} \in A, \mathbf{y} \in B$, $\mathbf{z} = \mathbf{y} - \mathbf{x} \in S$ and

$$f(\mathbf{z}) = f(\mathbf{y}) - f(\mathbf{x}) \geq c > 0$$

or

$$f(\mathbf{x}) + c \leq f(\mathbf{y})$$

which implies that

$$\sup_{\mathbf{x} \in A} f(\mathbf{x}) + c \leq \inf_{\mathbf{y} \in B} f(\mathbf{y})$$

and

$$\sup_{\mathbf{x} \in A} f(\mathbf{x}) < \inf_{\mathbf{y} \in B} f(\mathbf{y})$$

3.195 No. See Figure 3.2.

3.196 1. Assume that there exists a convex neighborhood $U \ni \mathbf{0}$ such that

$$(A + U) \cap B = \emptyset$$

Then $(A + U)$ is convex and $A \subset \text{int } (A + U) \neq \emptyset$ and $\text{int } (A + U) \cap B = \emptyset$. By Corollary 3.2.1, there exists continuous linear functional such that

$$f(\mathbf{x} + \mathbf{u}) \leq f(\mathbf{y}) \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{u} \in U, \mathbf{y} \in B$$

Since $f(U)$ is an open interval containing 0, there exists some $\mathbf{u}_0$ with $f(\mathbf{u}_0) = \epsilon > 0$.

$$f(\mathbf{x}) + \epsilon \leq f(\mathbf{y}) \qquad \text{for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

which implies that

$$\sup_{\mathbf{x} \in A} f(\mathbf{x}) < \inf_{\mathbf{y} \in B} f(\mathbf{y})$$

Conversely, assume that A and B can be strongly separated. That is, there exists a continuous linear functional $f \in X^*$ and number $\epsilon > 0$ such that

$$f(\mathbf{x}) \leq c - \epsilon < c + \epsilon \leq f(\mathbf{y}) \text{ for every } \ \mathbf{x} \in A, \mathbf{y} \in B$$

Let $U = \{ x \in X : |f(x)| < \epsilon \}$. U is a convex neighborhood of 0 such that $(A + U) \cap B = \emptyset$.

2. Let A and B be nonempty, disjoint, convex subsets in a normed linear space X with A compact and B closed. By Exercise 1.208, there exists a convex neighborhood $U \ni \mathbf{0}$ such that $(A+U) \cap B = \emptyset$. By the previous part, A and B can be strongly separated.

3.197 Assume $\rho(A,B) = \inf\{\, \|\mathbf{x}-\mathbf{y}\| : \mathbf{x} \in A, \mathbf{y} \in B \,\} = 2\epsilon > 0$. Let $U = B_\epsilon(\mathbf{0})$ be the open ball around $\mathbf{0}$ of radius ϵ. For every $\mathbf{x} \in A, \mathbf{u} \in U, \mathbf{y} \in B$

$$\|\mathbf{x} + (-\mathbf{u}) - \mathbf{y}\| = \|\mathbf{x} - \mathbf{y} - \mathbf{u}\| \geq \|\mathbf{x} - \mathbf{y}\| - \|\mathbf{u}\|$$

so that

$$\begin{aligned}
\rho(A+U,B) = \inf_{\mathbf{x},\mathbf{u},\mathbf{y}} \|\mathbf{x} + (-\mathbf{u}) - \mathbf{y}\| &\geq \inf_{\mathbf{x},\mathbf{u},\mathbf{y}} (\|\mathbf{x} - \mathbf{y}\| - \|\mathbf{u}\|) \\
&\geq \inf_{\mathbf{x},\mathbf{y}} \|\mathbf{x} - \mathbf{y}\| - \sup_{\mathbf{u}} \|\mathbf{u}\|) \\
&= 2\epsilon - \epsilon \\
&= \epsilon > 0
\end{aligned}$$

Therefore $(A+U) \cap B = \emptyset$ and so A and B can be strongly separated.

Conversely, assume that A and B can be strongly separated, so that there exists a convex neighborhood U of $\mathbf{0}$ such that $(A+U) \cap B = \emptyset$. Therefore, there exists $\epsilon > 0$ such that $B_\epsilon(\mathbf{0}) \subseteq U$ and

$$A + B_\epsilon \cap B = \emptyset$$

This implies that

$$\rho(A,B) = \inf\{\, \|\mathbf{x}-\mathbf{y}\| : \mathbf{x} \in A, \mathbf{y} \in B \,\} > \epsilon > 0$$

3.198 Take $A = \{\mathbf{y}\}$ and $B = M$ in Proposition 3.14. There exists $f \in X^*$ such that

$$f(\mathbf{y}) < c \leq f(\mathbf{x}) \qquad \text{for every } \mathbf{x} \in M$$

By Corollary 3.2.3, $c = 0$.

3.199 1. Consider the set

$$Z = \{\, f(x), -g_1(x), -g_2(x), \dots, -g_m(x) : x \in X \,\}$$

Z is the image of a linear mapping from X to $Y = \Re^{m+1}$ and hence is a subspace of $\Re^{m+1}$.

2. By hypothesis, the point $\mathbf{e}^0 = (1,0,0,\dots,0) \in \Re^{m+1}$ does not belong to Z. Otherwise, we have an $x \in X$ such that $g_i(x) = 0$ for every i but $f(x) = 1$.

3. By the previous exercise, there exists a linear functional $\varphi \in Y^*$ such that

$$\begin{aligned}
\varphi(\mathbf{e}^0) &> 0 \\
\varphi(\mathbf{z}) &= 0 \qquad \text{for every } \mathbf{z} \in Z
\end{aligned}$$

4. In other words, there exists a vector $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in Y = \Re^{(m+1)*}$ such that

$$\lambda \mathbf{e}^0 > 0 \tag{3.23}$$
$$\lambda \mathbf{z} = 0 \qquad \text{for every } \mathbf{z} \in Z \tag{3.24}$$

Equation (3.23) states that

$$\lambda \mathbf{z} = \lambda_0 \mathbf{z}_0 + \lambda_1 \mathbf{z}_1 + \cdots + \lambda_m \mathbf{z}_m = 0 \qquad \text{for every } \mathbf{z} \in Z$$

That is, for every $x \in X$,

$$\lambda_0 f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \lambda_2 g_2(\mathbf{x}) - \ldots - \lambda_m g_m(\mathbf{x}) = 0$$

5. Inequality (3.22) establishes that $\lambda_0 > 0$. Without loss of generality we can normalize so that $\lambda_0 = 1$.
6. Therefore

$$f(x) = \sum_{i=1}^m \lambda_i g_i(x)$$

3.200 For every $\mathbf{x} \in S$, $g_j(\mathbf{x}) = 0, j = 1, 2 \ldots m$ and therefore

$$f(\mathbf{x}) = \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) = 0$$

3.201 The set

$$Z = \{\, g_1(x), g_2(x), \ldots, g_m(x) : x \in X \,\}$$

is a closed subspace in $\Re^m$. If the system is inconsistent, $\mathbf{c} = (c_1, c_2, \ldots, c_m) \notin Z$. By Exercise 3.198, there exists a linear functional φ on $\Re^m$ such

$$\varphi(\mathbf{z}) = 0 \text{ for every } \mathbf{z} \in Z$$
$$\varphi(\mathbf{c}) > 0$$

That is, there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\sum_{j=1}^m \lambda_j g_j(\mathbf{x}) = 0$$

and

$$\sum_{j=1}^m \lambda_j c_j > 0$$

which contradicts the hypothesis

$$\sum_{j=1}^m \lambda_j g_j = 0 \implies \sum_{j=1}^m \lambda_j c_j = 0$$

Conversely, if for some $\mathbf{x} \in X$

$$g_j(\mathbf{x}) = c_j \qquad j = 1, 2, \ldots, m$$

then

$$\sum_{j=1}^m \lambda_j g_j(\mathbf{x}) = \sum_{j=1}^m \lambda_j c_j$$

and

$$\sum_{j=1}^m \lambda_j g_j = 0 \implies \sum_{j=1}^m \lambda_j c_j = 0$$

3.202 The set $\hat{K} = \{ \mathbf{x} \in K : \|\mathbf{x}\|_1 = 1 \}$ is

- compact (the unit ball is compact if and only if X is finite-dimensional)
- convex (which is why we need the 1 norm)

By Proposition 3.14, there exists a linear functional $f \in X^*$ such that

$$\begin{aligned} f(\hat{\mathbf{x}}) &> 0 \qquad \text{for every } \hat{\mathbf{x}} \in \hat{K} \\ f(\mathbf{x}) &= 0 \qquad \text{for every } \mathbf{x} \in M \end{aligned}$$

For any $\mathbf{x} \in K, \mathbf{x} \neq \mathbf{0}$, define $\hat{\mathbf{x}} = \mathbf{x} / \|\mathbf{x}\|_1 \in \hat{K}$. Then

$$f(\mathbf{x}) = f(\|\mathbf{x}\|_1 \hat{\mathbf{x}}) = \|\mathbf{x}\|_1 f(\hat{\mathbf{x}}) > 0$$

3.203 1. Let

$$\begin{aligned} A &= \{ (\mathbf{x}, y) : y \geq g(\mathbf{x}), \mathbf{x} \in X \} \\ B &= \{ (\mathbf{x}, y) : y = f_0(\mathbf{x}), \mathbf{x} \in Z \} \end{aligned}$$

A is the epigraph of a convex functional and hence convex. B is a subspace of $Y = X \times \Re$ and also convex.

2. Since g is convex, $\operatorname{int} A \neq \emptyset$. Furthermore

$$f_0(\mathbf{x}) \leq g(\mathbf{x}) \implies \operatorname{int} A \cap B = \emptyset$$

3. By Exercise 3.2.3, there exists linear functional $\varphi \in Y^*$ such that

$$\begin{aligned} \varphi(\mathbf{x}, y) &\geq 0 \qquad \text{for every } (\mathbf{x}, y) \in A \\ \varphi(\mathbf{x}, y) &= 0 \qquad \text{for every } (\mathbf{x}, y) \in B \end{aligned}$$

There exists y such that $y > g(\mathbf{0})$ and therefore $(\mathbf{0}, y) \in \operatorname{int} A$ and $\varphi(\mathbf{0}, y) > 0$. Therefore

$$\varphi(\mathbf{0}, 1) = \frac{1}{y} \varphi(\mathbf{0}, 1) > 0$$

4. Let $f \in X^*$ be defined by

$$f(\mathbf{x}) = -\frac{1}{c} \varphi(\mathbf{x}, 0)$$

where $c = \varphi(\mathbf{0}, 1)$. Since

$$\begin{aligned} \varphi(\mathbf{x}, 0) &= \varphi(\mathbf{x}, y) - \varphi(\mathbf{0}, y) \\ &= \varphi(\mathbf{x}, y) - cy \end{aligned}$$

$$f(\mathbf{x}) = -\frac{1}{c}(\varphi(\mathbf{x}, y) - cy) = -\frac{1}{c} \varphi(\mathbf{x}, y) + y$$

for every $y \in \Re$

5. For every $\mathbf{x} \in Z$

$$\begin{aligned} f(\mathbf{x}) &= -\frac{1}{c} \varphi(\mathbf{x}, f_0(\mathbf{x})) + f_0(\mathbf{x}) \\ &= f_0(\mathbf{x}) \end{aligned}$$

since $\varphi(\mathbf{x}, f_0(\mathbf{x})) = 0$ for every $\mathbf{x} \in Z$. Thus f is an extension of f_0.

6. For any $\mathbf{x} \in X$, let $y = g(\mathbf{x})$. Then $(\mathbf{x}, y) \in A$ and $\varphi(\mathbf{x}, y) \geq 0$. Therefore

$$\begin{aligned} f(\mathbf{x}) &= -\frac{1}{c}\varphi(\mathbf{x}, y) + y \\ &= -\frac{1}{c}\varphi(\mathbf{x}, y) + g(\mathbf{x}) \\ &\leq g(\mathbf{x}) \end{aligned}$$

Therefore f is bounded by g as required.

3.204 Let $g \in X^*$ be defined by

$$g(\mathbf{x}) = \|f_0\|_Z \|\mathbf{x}\|$$

Then $f_0(\mathbf{x}) \leq g(\mathbf{x})$ for all $\mathbf{x} \in Z$. By the Hahn-Banach theorem (Exercise 3.15), there exists an extension $f \in X^*$ such that

$$f(\mathbf{x}) \leq g(\mathbf{x}) = \|f_0\|_Z \|\mathbf{x}\|$$

Therefore

$$\|f\|_X = \sup_{\|\mathbf{x}\|=1} \|f(\mathbf{x})\| = \|f_0\|_Z$$

3.205 If $\mathbf{x}_0 = \mathbf{0}$, any bounded linear functional will do. Therefore, assume $\mathbf{x}_0 \neq \mathbf{0}$. On the subspace $\text{lin}\,\{\mathbf{x}_0\} = \{\alpha\mathbf{x}_0 : \alpha \in \Re\}$, define the function

$$f_0(\alpha\mathbf{x}_0) = \alpha \|\mathbf{x}_0\|$$

f_0 is a bounded linear functional on $\text{lin}\,\{\mathbf{x}_0\}$ with norm 1. By the previous part, f_0 can be extended to a bounded linear functional $f \in X^*$ with the same norm, that is $\|f\| = 1$ and $f(\mathbf{x}_0) = \|\mathbf{x}_0\|$.

3.206 Since $\mathbf{x}_1 \neq \mathbf{x}_2$, $\mathbf{x}_1 - \mathbf{x}_2 \neq \mathbf{0}$. There exists a bounded linear functional such that

$$f(\mathbf{x}_1 - \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\| \neq 0$$

so that

$$f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$$

3.207 1. • $\mathfrak{F}$ is a complete lattice (Exercise 1.179).

• The intersection of any chain is
  – nonempty (since S is compact)
  – a face (Exercise 1.179)

  Hence every chain has a minimal element.

• By Zorn's lemma (Remark 1.5), $\mathfrak{F}$ has a minimal element F_0.

2. Assume to the contrary that F_0 contains two distinct elements $\mathbf{x}_1, \mathbf{x}_2$. Then (Exercise 3.206) there exists a continuous linear functional $f \in X^*$ such that

$$f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$$

Let c be in the minimum value of $f(\mathbf{x})$ on F_0 and let F_1 be the set on which it attains this minimum. (Since F_0 is compact, c is well-defined and F_1 is nonempty. That is

$$\begin{aligned} c &= \min\{\, f(x) : \mathbf{x} \in F_0 \,\} \\ F_1 &= \{\, \mathbf{x} \in S : f(\mathbf{x}) = c \,\} \end{aligned}$$

Now $F_1 \subset F_0$ since $f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$.

To show that F_1 is a face of F_0, assume that $\alpha\mathbf{x}+(1-\alpha)\mathbf{y} \in F_1$ for some $\mathbf{x}, \mathbf{y} \in F_0$. Then $c = f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) = c$. Since $\mathbf{x}, \mathbf{y} \in F_0$, this implies that $f(\mathbf{x}) = f(\mathbf{y}) = c$ so that $\mathbf{x}, \mathbf{y} \in F_1$. Therefore F_1 is a face.

We have shown that, if F_0 contains two distinct elements, there exists a smaller face $F_1 \subset F_0$, contradicting the minimality of F_0. We conclude that F_0 comprises a single element $\mathbf{x}_0$.

3. $F_0 = \{\mathbf{x}_0\}$ which is an extreme point of S.

3.208 Let $H = H_f(c)$ be a supporting hyperplane to S. Without loss of generality assume

$$f(\mathbf{x}) \leq c \text{ for every } \mathbf{x} \in S \tag{3.25}$$

and there exists some $\mathbf{x}^* \in S$ such that

$$f(\mathbf{x}^*) = c$$

That is f is maximized at $\mathbf{x}^*$.

Version 1 By the previous exercise, f achieves its maximum at an extreme point. That is, there exists an extreme point $\mathbf{x}_0 \in S$ such that

$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \text{ for every } \mathbf{x} \in S$$

In particular, $f(\mathbf{x}_0) \geq f(\mathbf{x}^*) = c$. But (3.25) implies $f(\mathbf{x}_0) \leq c$. Therefore, we conclude that $f(\mathbf{x}_0) = c$ and therefore $\mathbf{x}_0 \in H$.

Version 2 The set $H \cap S$ is a nonempty, compact, convex subset of a linear space. Hence, by Exercise 3.207, $H \cap S$ contains an extreme point, say $\mathbf{x}_0$. We show that $\mathbf{x}_0$ is an extreme point of S.

Assume not, that is assume that there exists $\mathbf{x}_1, \mathbf{x}_2 \in S$ such that $\mathbf{x}_0 = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ for some $\alpha \in (0, 1)$. Since $\mathbf{x}_0$ *is* an extreme point of $H \cap S$, at least one of the points $\mathbf{x}_1, \mathbf{x}_2$ must lie outside H. Assume $\mathbf{x}_1 \notin H$ which implies that $f(\mathbf{x}_1) < c$. Since $f(\mathbf{x}_2) \leq c$

$$f(\mathbf{x}_0) = \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) < c \tag{3.26}$$

However, since $\mathbf{x}_0 \in H \cap S$, we must have

$$f(\mathbf{x}_0) = c$$

which contradicts (3.26).

Therefore $\mathbf{x}_0$ is an extreme point of S. In fact, we have shown that every extreme point of $H \cap S$ must be an extreme point of S.

3.209 Let $\hat{S}$ denote the closed, convex hull of the extreme points of S. (The closed, convex hull of a set is simply the closure of the convex hull.) Clearly $\hat{S} \subset S$ and it remains to show that $\hat{S}$ contains all of S.

Assume not. That is, assume $\hat{S} \subsetneq S$ and let $\mathbf{x}_0 \in S \setminus \hat{S}$. By the Strong Separation Theorem, there exists a linear functional $f \in X^*$ such that

$$f(\mathbf{x}_0) > f(\mathbf{x}) \text{ for every } \mathbf{x} \in \hat{S} \tag{3.27}$$

On the other hand, by Exercise 3.16, f attains its maximum at an extreme point of S. That is, there exists $\mathbf{x}_1 \in \hat{S}$ such that

$$f(\mathbf{x}_1) \geq f(\mathbf{x}) \quad \text{for every} \quad \mathbf{x} \in S$$

In particular

$$f(\mathbf{x}_1) \geq f(\mathbf{x}_0)$$

since $\mathbf{x}_0 \in \hat{S} \subset S$. This contradicts (3.27) since $\mathbf{x}_1 \in \hat{S}$.

Thus our assumption that $S \subsetneq \hat{S}$ yields a contradiction. We conclude that

$$S = \hat{S}$$

3.210 1. (a) P is compact and convex, since it is the product of compact, convex sets (Proposition 1.2, Exercise 1.165).

(b) Since $\mathbf{x} \in \sum_{i=1}^n \text{conv } S_i$, there exist $\mathbf{x}_i \in \text{conv } S_i$ such that $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i$. $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \in P(\mathbf{x})$ so that $P(\mathbf{x}) \neq \emptyset$.

(c) By the Krein-Millman theorem (or Exercise 3.207), $P(\mathbf{x})$ has an extreme point $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)$ such that

- $\mathbf{z}_i \in \text{conv } S_i$ for every i
- $\sum_{i=1}^n \mathbf{z}_i = \mathbf{x}$.

since $\mathbf{z} \in P(\mathbf{x})$.

2. (a) Exercise 1.176

(b) Since $l > m = \dim X$, the vectors $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_l$ are linearly dependent (Exercise 1.143). Consequently, there exists numbers $\alpha'_1, \alpha'_2, \dots, \alpha'_l$, not all zero, such that

$$\alpha'_1 \mathbf{y}_1 + \alpha'_2 \mathbf{y}_2 + \dots + \alpha'_l \mathbf{y}_l = 0$$

(Exercise 1.133). Let

$$\alpha_i = \frac{\alpha'_i}{\max_i |\alpha_i|}$$

Then $|\alpha_i| \leq 1$ for every i and

$$\alpha_1 \mathbf{y}_1 + \alpha_2 \mathbf{y}_2 + \dots + \alpha_l \mathbf{y}_l = 0$$

(c) Since $|\alpha_i| \leq 1$, $\mathbf{z}_i + \alpha_i \mathbf{y}_i \in \text{conv } S_i$ for every $i = 1, 2, \dots, l$. Furthermore

$$\sum_{i=1}^n \mathbf{z}_i^+ = \sum_{i=1}^n \mathbf{z}_i + \sum_{i=1}^l \alpha_i \mathbf{y}_i = \sum_{i=1}^n \mathbf{z}_i = \mathbf{x}$$

Therefore, $\mathbf{z}^+ \in P(\mathbf{x})$. Similarly, $\mathbf{z}^- \in P(\mathbf{x})$.

(d) By direct computation

$$\mathbf{z} = \frac{1}{2}\mathbf{z}^+ + \frac{1}{2}\mathbf{z}^-$$

which implies that $\mathbf{z}$ is not an extreme point of $P(\mathbf{x})$, contrary to our assumption. This establishes that at least $n - m$ $\mathbf{z}_i$ are extreme points of the corresponding conv S_i.

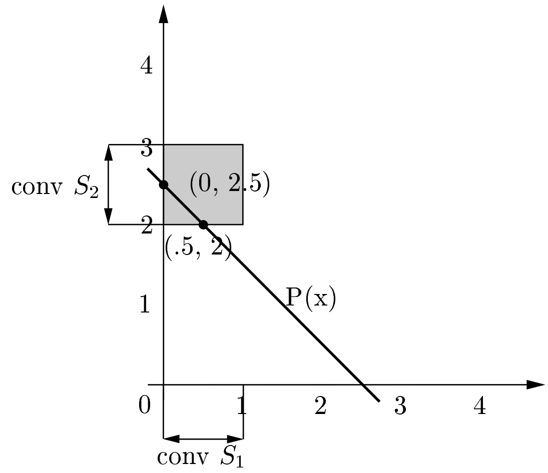

Figure 3.3: Illustrating the proof of the Shapley Folkman theorem.

3. Every extreme point of conv S_i is an element of S_i.

3.211 See Figure 3.3.

3.212 Let $\{S_1, S_2, \dots, S_n\}$ be a collection of nonempty subsets of an m-dimensional linear space and let $\mathbf{x} \in \text{conv} \sum_{i=1}^n S_i = \sum_{i=1}^n \text{conv } S_i$. That is, there exists $\mathbf{x}_i \in$ conv S_i such that $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i$. By Carathéodory's theorem, there exists for every $\mathbf{x}_i$ a finite number of points $\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{il_i}$ such that $\mathbf{x}_i \in \text{conv } \{\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{il_i}\}$.

For every $i = 1, 2, \dots, n$, let

$$\tilde{S}_i = \{\, \mathbf{x}_{ij} : j = 1, 2, \dots, l_i \,\}$$

Then

$$\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i, \qquad \mathbf{x}_i \in \text{conv } \tilde{S}_i$$

That is, $\mathbf{x} \in \sum \text{conv } \tilde{S}_i = \text{conv } \sum \tilde{S}_i$. Moreover, the sets S_i are compact (in fact finite). By the previous exercise, there exists n points $\mathbf{z}_i \in \tilde{S}_i$ such that

$$\mathbf{x} = \sum_{i=1}^n \mathbf{z}_i, \qquad \mathbf{z}_i \in \text{conv } \tilde{S}_i$$

and moreover $\mathbf{z}_i \in \tilde{S}_i \subseteq S_i$ for at least $n - m$ indices i.

3.213 Let S be a closed convex set in a normed linear space. Clearly, S is contained in the intersection of all the closed halfspaces which contain S.

For any $\mathbf{y} \notin S$, there exists a hyperplane which strongly separates $\{\mathbf{y}\}$ and S. One of its closed halfspaces contains S but not $\mathbf{y}$. Consequently, $\mathbf{y}$ does not belong to the intersection of all the closed halfspaces containing S.

3.214 1. Since $V^*(y)$ is the intersection of closed, convex sets, it is closed and convex. Assume $\mathbf{x}$ is feasible, that is $\mathbf{x} \in V(y)$. Then $\mathbf{w}^T\mathbf{x} \leq (c\mathbf{w}, y)$ and $\mathbf{x} \in V^*(y)$. That is, $V(y) \subseteq V^*(y)$.

2. Assume $V(y)$ is convex. For any $\mathbf{x}_0 \notin V(y)$ there exists $\mathbf{w}$ such that

$$\mathbf{w}^T\mathbf{x}_0 < \inf_{\mathbf{x} \in V(y)} \mathbf{w}^T\mathbf{x} = c(\mathbf{w}, y)$$

by the Strong Separation Theorem. Monotonicity ensures that $\mathbf{w} \geq \mathbf{0}$ and hence $\mathbf{x}_0 \notin V^*(y)$.

3.215 Assume $\mathbf{x} \in V(y) = V^*(y)$. That is

$$\mathbf{w}^T\mathbf{x} \geq y\hat{c}(\mathbf{w}) \text{ for every } \mathbf{x}$$

Therefore, for any $t \in \Re_+$

$$t\mathbf{w}^T\mathbf{x} \geq tyc(\mathbf{w}) \text{ for every } \mathbf{x}$$

which implies that $t\mathbf{x} \in V^*(y) = V(y)$.

3.216 A polyhedron

$$\begin{aligned} S &= \{\, x \in X : g_i(x) \leq c_i,\ i = 1, 2, \ldots, m \,\} \\ &= \bigcap_{i=1}^{m} \{\, \mathbf{x} \in X : g_i(\mathbf{x}) \leq c_i \,\} \end{aligned}$$

is the intersection of a finite number of closed convex sets.

3.217 Each row $\mathbf{a}^i = (a_{i1}, a_{i2}, \ldots a_{in})$ of A defines a linear functional $g_i(\mathbf{x}) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$ on $\Re^n$. The set S of solutions to $A\mathbf{x} \leq \mathbf{c}$ is

$$S = \{\, x \in X : g_i(\mathbf{x}) \leq c_i,\ i = 1, 2, \ldots, m \,\}$$

is a polyhedron.

3.218 For simplicity, we assume that the game is superadditive, so that $w(i) \geq 0$ for every i. Consequently, in every core allocation $\mathbf{x}$, $0 \leq x_i \leq w(N)$ and

$$\text{core} \subseteq [0, w(N)] \times [0, w(N)] \times \cdots \times [0, w(N)] \subset \Re^n$$

Thus, the core is bounded. Since it is the intersection of closed halfspaces, the core is also closed. By Proposition 1.1, the core is compact.

3.219 **polytope $\implies$ polyhedron** Assume that P is a polytope generated by the points $\{\, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \,\}$ and let $F_1, F_2, \ldots, F_k$ denote the proper faces of P. For each $i = 1, 2, \ldots, k$, let H_i denote the hyperplane containing F_i so that $F_i = P \cap H_i$. For every such hyperplane, there exists a nonzero linear functional g_i and constant c_i such that $g_i(\mathbf{x}) = c_i$ for every $\mathbf{x} \in H_i$. Furthermore, every such hyperplane is a bounding hyperplane of P. Without loss of generality, we can assume that $g_i(\mathbf{x}) \leq c$ for every $\mathbf{x} \in P$. Let

$$S = \{\, \mathbf{x} \in X : g_i(\mathbf{x}) \leq c_i,\ i = 1, 2, \ldots, m \,\}$$

Clearly $P \subseteq S$. To show that $S \subseteq P$, assume not. That is, assume that there exists $\mathbf{y} \in S \setminus P$ and let $\mathbf{x} \in \text{ri } P$. (ri P is nonempty by exercise 1.229). Since P is closed (Exercise 1.227), there exists a some α such that $\bar{\mathbf{x}} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ belongs to the relative boundary of P, and there exists some i such that $\bar{\mathbf{x}} \in F_i \subseteq H_i$.

Let $H_i^+ = \{\, \mathbf{x} \in X : g_i(\mathbf{x}) \leq c_i \,\}$ denote the closed half-space bounded by H_i and containing P. H_i is a face of H_i^+ containing $\bar{\mathbf{x}} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, which implies that $\mathbf{x}, \mathbf{y} \in H_i$. This in turn implies that $\mathbf{x} \in F_i$, which contradicts the assumption that $\mathbf{x} \in \text{ri } P$. We conclude that $S = P$.

polyhedron $\Longrightarrow$ polytope Conversely, assume S is a nonempty compact polyhedral set in a normed linear space. Then, there exist linear functionals $g_1, g_2, \dots, g_m$ in X^* and numbers $c_1, c_2, \dots, c_m$ such that $S = \{ \mathbf{x} \in X : g_i(\mathbf{x}) \leq c_i,\ i = 1, 2, \dots, m \}$. We show that S has a finite number of extreme points. Let n denote the dimension of S. If $n = 1$, S is either a single point or closed line segment (since S is compact), and therefore has a finite number of extreme points (that is, 1 or 2).

Now assume that every compact polyhedral set of dimension $n-1$ has a finite number of extreme points. Let $H_i,\ i = i = 1, 2, \dots, m$ denote the hyperplanes associated with the linear functionals g_i defining S (Exercise 3.49). Let $\mathbf{x}$ be an extreme point of S. Then S is a boundary point of S (Exercise 1.220) and therefore belongs to some H_j. We claim that $\mathbf{x}$ is also an extreme point of the set $S \cap H_j$. To see this, assume otherwise. That is, assume that $\mathbf{x}$ is not an extreme point of $S \cap H_j$. Then, there exists $\mathbf{x}_1, \mathbf{x}_2 \in S \cap H_j$ such that $\mathbf{x} = \alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$. But then $\mathbf{x}_1, \mathbf{x}_2 \in S$ and $\mathbf{x}$ is not an extreme point of S. Therefore, every extreme point of S is an extreme point of some $S \cap H_i$, which is a compact polyhedral set of dimension $n-1$. By hypothesis, each $S \cap H_i$ has a finite number of extreme points. Since there are only m such hyperplanes H_i, S has a finite number of extreme points.

By the Krein-Milman theorem (Exercise 3.209), S is the closed convex hull of its extreme points. Since there are only finite extreme points, S is a polytope.

3.220 1. Let $f, g \in S^*$ so that $f(\mathbf{x}) \leq 0$ and $g(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in S$. For every $\alpha, \beta \geq 0$

$$\alpha f(\mathbf{x}) + \beta f(\mathbf{x}) \leq 0$$

for every $x \in S$. This shows that $\alpha f + \beta g \in S^*$. S^* is a convex cone.

To show that S^* is closed, let f be the limit of a sequence (f_n) of functionals in S^*. Then, for every $\mathbf{x} \in S$,

$$f_n(\mathbf{x}) \leq 0$$

so that

$$f(\mathbf{x}) = \lim f_n(\mathbf{x}) \leq 0$$

2. Let $\mathbf{x}, \mathbf{y} \in S^{**}$. Then, for every $f \in S^*$

$$f(\mathbf{x}) \leq 0 \text{ and } f(\mathbf{y}) \leq 0$$

and therefore

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \leq 0$$

for every $\alpha, \beta \geq 0$. There $\alpha\mathbf{x} + \beta\mathbf{y} \in S^{**}$. S^{**} is a convex cone.

To show that S^{**} is closed, let $\mathbf{x}^n$ be a sequence of points in S^{**} converging to x. For every $n = 1, 2, \dots$

$$f(\mathbf{x}^n) \leq 0 \text{ for every } f \in S^*$$

By continuity

$$f(\mathbf{x}) = \lim f(\mathbf{x}^n) \leq 0 \text{ for every } f \in S^*$$

Consequently $\mathbf{x} \in S^{**}$ which is therefore closed.

3. Let $\mathbf{x} \in S$. Then $f(\mathbf{x}) \leq 0$ for every $f \in S^*$ so that $\mathbf{x} \in S^{**}$.
4. Exercise 1.79.

3.221 Let $f \in S_2^*$. Then $f(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in S_2$. *A fortiori*, since $S_1 \subseteq S_2$, $f(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in S_1$. Therefore $f \in S_1^*$.

3.222 Exercise 3.220 showed that $S \subseteq S^{**}$. To show the converse, let $\mathbf{y} \notin S$. By Proposition 3.14, there exists some $f \in X^*$ and c such that

$$\begin{aligned} f(\mathbf{y}) &> c \\ f(\mathbf{x}) &< c \qquad \text{for every } \mathbf{x} \in S \end{aligned}$$

Since S is a cone, $\mathbf{0} \in S$ and $f(\mathbf{0}) = 0 < c$. Since $\alpha S = S$ for every $\alpha > 0$ then

$$f(\mathbf{x}) < 0 \qquad \text{for every } \mathbf{x} \in S$$

so that $f \in S^*$. $f(\mathbf{y}) > 0$, $\mathbf{y} \notin S^{**}$. That is

$$\mathbf{y} \notin S \implies \mathbf{y} \notin S^{**}$$

from which we conclude that $S^{**} \subseteq S$.

3.223 Let

$$\begin{aligned} K &= \text{cone}\,\{g_1, g_2, \ldots, g_m\} \\ &= \{\, g \in X^* : g = \sum_{j=1}^m \lambda_j g_j, \lambda_j \geq 0 \,\} \end{aligned}$$

be the set of all nonnegative linear combinations of the linear functionals g_j. K is a closed convex cone.

Suppose that $f \notin \text{cone}\,\{g_1, g_2, \ldots, g_m\}$, that is assume that $f \notin K$. Then $\{f\}$ is a compact convex set disjoint from K. By Proposition 3.14, there exists a continuous linear functional φ and number c such that

$$\sup_{g \in K} \varphi(g) < c < \varphi(f)$$

Since $\mathbf{0} \in K$, $c \geq 0$ and so $\varphi(f) > 0$. Further, for every $g \in G$

$$\begin{aligned} \varphi(g) &= \varphi(\sum_{j=1}^m \lambda_j g_j) \\ &= \sum_{j=1}^m \lambda_j \varphi(g_j) < c \text{ for every } \lambda_j \geq 0 \end{aligned}$$

Since λ_j can be made arbitrarily large, this last inequality implies that

$$\varphi(g_j) \leq 0 \qquad j = 1, 2, \ldots, m$$

By the Riesz representation theorem (Exercise 3.75), there exists $\mathbf{x} \in X$

$$\varphi(g_j) = g_j(\mathbf{x}) \text{ and } \varphi(f) = f(\mathbf{x})$$

Since

$$\varphi(g_j) = g_j(\mathbf{x}) \leq 0$$

$\mathbf{x} \in S$. By hypothesis

$$f(\mathbf{x}) = \varphi(f) \leq 0$$

contradicting the conclusion that $\varphi(f) > 0$. This contradiction establishes that $f \in K$, that is

$$f(x) = \sum_{j=1}^{m} \lambda_j g_j(x), \qquad \lambda_j \geq 0$$

3.224 Let $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^m$ denote the rows of A and define the linear functional $f, g_1, g_2, \dots, g_m$ by

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{cx} \\ g_j(\mathbf{x}) &= \mathbf{a}^j\mathbf{x} \ j = 1, 2, \dots, m \end{aligned}$$

Assume $\mathbf{cx} \leq 0$ for every $\mathbf{x}$ satisfying $A\mathbf{x} \leq \mathbf{0}$, that is $f(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in S$ where

$$S = \{\, \mathbf{x} \in X : g_j(\mathbf{x}) \leq 0, \ j = 1, 2, \dots, m \,\}$$

By Proposition 3.18, there exists $\mathbf{y} \in \Re^m_+$ such that

$$f(\mathbf{x}) = \sum_{j=1}^{m} y_j g_j(\mathbf{x})$$

or

$$\mathbf{c} = \sum_{j=1}^{m} y_j \mathbf{a}^j = A^T\mathbf{y}$$

Conversely, assume that

$$\mathbf{c} = A^T\mathbf{y} = \sum_{j=1}^{m} y_j \mathbf{a}^j$$

Then

$$A\mathbf{x} \leq \mathbf{0} \implies \mathbf{a}^j\mathbf{x} \leq 0 \ \text{ for every } \ j \implies \mathbf{cx} \leq 0$$

3.225 Let $N = \Re^n_+$ denote the positive orthant of $\Re^n$. N is a convex set (indeed cone) with a nonempty interior. By Corollary 3.2.1, there exists a hyperplane $H_{\mathbf{p}}(c)$ such that

$$\mathbf{p}^T\mathbf{x} \leq c \leq \mathbf{py} \quad \text{for every } \ \mathbf{x} \in S, \mathbf{y} \in N$$

Since $\mathbf{0} \in N$

$$\mathbf{p0} = 0 \geq c$$

which implies that $c \leq 0$ and

$$\mathbf{p}^T\mathbf{x} \leq c \leq 0 \quad \text{for every } \ x \in S$$

To show that $\mathbf{p}$ is nonnegative, let $\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^n$ denote the standard basis for $\Re^n$. Each $\mathbf{e}^i$ belongs to N so that

$$\mathbf{pe}^i = p_i \geq 0 \quad \text{for every } \ i$$

3.226 Assume $\mathbf{y}^*$ is an efficient production plan in Y and let $S = Y - y^*$. S is convex. We claim that $S \cap \Re^n_{++} = \emptyset$. Otherwise, if there exists some $\mathbf{z} \in S \cap \Re^n_{++}$, let $\mathbf{y}' = \mathbf{y}^* + \mathbf{z}$

- $\mathbf{z} \in S$ implies $\mathbf{y}' \in Y$ while
- $\mathbf{z} \in \Re^n_{++}$ implies $\mathbf{y}' > \mathbf{y}^*$

contradicting the efficiency of $\mathbf{y}^*$. Therefore, S is a convex set which contains no interior points of the nonnegative orthant $\Re^n_+$. By Exercise 3.225, there exists a price system $\mathbf{p}$ such that

$$\mathbf{p}^T\mathbf{x} \leq 0 \text{ for every } \mathbf{x} \in S$$

Since $S = Y - y^*$, this implies

$$\mathbf{p}(\mathbf{y} - \mathbf{y}^*) \leq 0 \text{ for every } \mathbf{y} \in Y$$

or

$$\mathbf{p}\mathbf{y}^* \geq \mathbf{p}\mathbf{y} \text{ for every } \mathbf{y} \in Y$$

y^* maximizes the producer's profit at prices $\mathbf{p}$.

3.227 Consider the set $S^- = \{\, \mathbf{x} \in \Re^n : -\mathbf{x} \in S \,\}$.

$$S \cap \operatorname{int} \Re^n_- = \emptyset \implies S^- \cap \operatorname{int} \Re^n_+ = \emptyset$$

From the previous exercise, there exists a hyperplane with nonnegative normal $\mathbf{p} \gneq \mathbf{0}$ such that

$$\mathbf{p}^T\mathbf{x} \leq 0 \quad \text{for every } \mathbf{x} \in S^-$$

Since $\mathbf{p} \gneq \mathbf{0}$, this implies

$$\mathbf{p}^T\mathbf{x} \geq 0 \quad \text{for every } \mathbf{x} \in S$$

3.228 1. Suppose $\mathbf{x} \in \succsim(\mathbf{x}^*)$. Then, there exists an allocation $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ such that

$$\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i$$

where $\mathbf{x}_i \in \succsim(\mathbf{x}_i^*)$ for every $i = 1, 2, \dots, n$. Conversely, if $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ is an allocation with $\mathbf{x}_i \in \succsim(\mathbf{x}_i^*)$ for every $i = 1, 2, \dots, n$, then $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i \in \succsim(\mathbf{x}^*)$.

2. For every agent i, $\mathbf{x}_i^* \in \succsim(\mathbf{x}_i^*)$, which implies that

$$\mathbf{x}^* = \sum_{i=1}^n \mathbf{x}_i^* \in \succsim(\mathbf{x}^*)$$

and therefore

$$\mathbf{0} \in S = \succsim(\mathbf{x}^*) - \mathbf{x}^* \neq \emptyset$$

Since individual preferences are convex, $\succsim(\mathbf{x}_i^*)$ is convex for each i and therefore $S = \succsim(\mathbf{x}^*) - \mathbf{x}^* = \sum_i \succsim(\mathbf{x}_i^*) - \mathbf{x}^*$ is convex (Exercise 1.164).

Assume to the contrary that $S \cap \operatorname{int} \Re^l_- \neq \emptyset$. That is, there exists some $\mathbf{z} \in S$ with $\mathbf{z} < \mathbf{0}$. This implies that there exists some allocation $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ such that

$$\mathbf{z} = \sum_i \mathbf{x}_i - \mathbf{x}^* < \mathbf{0}$$

and $\mathbf{x}_i \succsim \mathbf{x}_i^*$ for every $i \in N$. Distribute $\mathbf{z}$ equally to all the consumers. That is, consider the allocation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{z}/n$$

By strict monotonicity, $\mathbf{y}_i \succ \mathbf{x}_i \succsim \mathbf{x}_i^*$ for every $i \in N$. Since

$$\sum_i \mathbf{y}_i = \sum_i \mathbf{x}_i + \mathbf{z} = \mathbf{x}^* = \sum_i \mathbf{x}_i^*$$

$(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$ is a reallocation of the original allocation $\underline{\mathbf{x}}^*$ which is strictly preferred by all consumers. This contradicts the assumed Pareto efficiency of $\underline{\mathbf{x}}^*$. We conclude that

$$S \cap \operatorname{int} \Re^l_- \neq \emptyset$$

3. Applying Exercise 3.227, there exists a hyperplane with nonnegative normal $\mathbf{p}^* \gneq \mathbf{0}$ such that

$$\mathbf{p}^*\mathbf{z} \geq \mathbf{0} \text{ for every } \mathbf{z} \in S$$

That is

$$\mathbf{p}^*(\mathbf{x} - \mathbf{x}^*) \geq \mathbf{0} \text{ or } \mathbf{p}^*\mathbf{x} \geq \mathbf{p}^*\mathbf{x}^* \text{ for every } \mathbf{x} \in \succsim(\mathbf{x}^*) \tag{3.28}$$

4. Consider any allocation which is strictly preferred to $\underline{\mathbf{x}}^*$ by consumer j, that is $\mathbf{x}_j \in \succ_j(\mathbf{x}_j^*)$. Construct another allocation $\underline{\mathbf{y}}$ by taking $\epsilon > 0$ of *each* commodity away from agent j and distributing amongst the other agents to give

$$\begin{aligned} \mathbf{y}_j &= (1-\epsilon)\mathbf{x}_j \\ \mathbf{y}_i &= \mathbf{x}_i^* + \frac{\epsilon}{n-1}\mathbf{x}_j, \qquad i \neq j \end{aligned}$$

By continuity, there exists some $\epsilon > 0$ such that $\mathbf{y}_j = (1-\epsilon)\mathbf{x}_j \succ_j \mathbf{x}_j^*$. By monotonicity, $\mathbf{y}_i \succ_i \mathbf{x}_i^*$ for every $i \neq j$. We have constructed an allocation $\underline{\mathbf{y}}$ which is strictly preferred to $\underline{\mathbf{x}}^*$ by all the agents, so that $\mathbf{y} = \sum_i \mathbf{y}_i \in \succsim(\mathbf{x}^*)$. (3.28) implies that

$$\mathbf{p}\mathbf{y} \geq \mathbf{p}\mathbf{x}^*$$

That is

$$\mathbf{p}\left((1-\epsilon)\mathbf{x}_j + \sum_{i \neq j}\left(\mathbf{x}_i^* + \frac{\epsilon}{n-1}\mathbf{x}_j\right)\right) = \mathbf{p}\left(\mathbf{x}_j + \sum_{i \neq j}\mathbf{x}_i^*\right) \geq \mathbf{p}\left(\mathbf{x}_j^* + \sum_{i \neq j}\mathbf{x}_i^*\right)$$

which implies that

$$\mathbf{p}\mathbf{x}_j \geq \mathbf{p}\mathbf{x}_j^* \qquad \text{for every } \mathbf{x}_j \in \succ(\mathbf{x}_j^*) \tag{3.29}$$

5. Trivially, $\underline{\mathbf{x}}^*$ is a feasible allocation with endowments $\mathbf{w}_i = \mathbf{x}_i^*$ and $m_i = \mathbf{p}^*\mathbf{w}_i = \mathbf{p}^*\mathbf{x}_i^*$. To show that $(\mathbf{p}^*, \underline{\mathbf{x}}^*)$ is a competitive equilibrium, we have to show that $\mathbf{x}_i^*$ is the best allocation in the budget set $X_i(\mathbf{p}, m_i)$ for each consumer i. Suppose to the contrary there exists some consumer j and allocation $\mathbf{y}_j$ such that $\mathbf{y}_j \succ \mathbf{x}_j$ and $\mathbf{p}\mathbf{y}_j \leq m_j = \mathbf{p}\mathbf{x}_j^*$. By continuity, there exists some $\alpha \in (0,1)$ such that $\alpha\mathbf{y}_j \succ_i \mathbf{x}_j^*$ and

$$\mathbf{p}(\alpha\mathbf{y}_j) = \alpha\mathbf{p}\mathbf{y}_j < \mathbf{p}\mathbf{y}_j \leq \mathbf{p}\mathbf{x}^*$$

contradicting (3.29). We conclude that

$$\mathbf{x}_i^* \succsim_i \mathbf{x}_i \text{ for every } \mathbf{x} \in X(\mathbf{p}^*, m_i)$$

for every consumer i. $(\mathbf{p}^*, \underline{\mathbf{x}}^*)$ is a competitive equilibrium.

3.229 By the previous exercise, there exists a price system $\mathbf{p}^*$ such that $\mathbf{x}_i^*$ is optimal for each consumer i in the budget set $X(\mathbf{p}^*, \mathbf{p}^*\mathbf{x}_i^*)$, that is

$$\mathbf{x}_i^* \succsim_i \mathbf{x}_i \text{ for every } \mathbf{x}_i \in X(\mathbf{p}^*, \mathbf{p}^*\mathbf{x}_i^*) \tag{3.30}$$

For each consumer, let t_i be the difference between her endowed wealth $\mathbf{p}^*\mathbf{w}_i$ and her required wealth $\mathbf{p}^*\mathbf{x}_i^*$. That is, define

$$t_i = \mathbf{p}^*\mathbf{x}_i^* - \mathbf{p}^*\mathbf{w}_i = \mathbf{p}^*(\mathbf{x}_i^* - \mathbf{w}_i)$$

Then

$$\mathbf{p}^*\mathbf{x}_i^* = \mathbf{p}^* + \mathbf{w}_i \tag{3.31}$$

By assumption $\underline{\mathbf{x}}^*$ is feasible, so that

$$\sum_i \mathbf{x}_i^* - \sum_i \mathbf{w}_i = \sum_i (\mathbf{x}_i^* - \mathbf{w}_i) = 0$$

so that

$$\sum_i t_i = \mathbf{p}^* \sum_i (\mathbf{x}_i^* - \mathbf{w}_i) = 0$$

Furthermore, for $m_i = p^*\mathbf{w}_i + t_i$, (3.31) implies

$$X(\mathbf{p}^*, m_i) = \{\, \mathbf{x}_i : \mathbf{p}^*\mathbf{x}_i \leq \mathbf{p}^*\mathbf{w}_i + t_i \,\} = \{\, \mathbf{x}_i : \mathbf{p}^*\mathbf{x}_i \leq \mathbf{p}^*\mathbf{x}_i^* \,\} = X(\mathbf{p}^*, \mathbf{p}^*\mathbf{x}_i^*)$$

for each consumer i. Using (3.30) we conclude that

$$\mathbf{x}_i^* \succsim_i \mathbf{x}_i \text{ for every } \mathbf{x}_i \in X(\mathbf{p}^*, m_i)$$

for every agent i. $(\mathbf{p}^*, \underline{\mathbf{x}}^*)$ is a competitive equilibrium where each consumer's after-tax wealth is

$$m_i = \mathbf{p}\mathbf{w}_i + t_i$$

3.230 Apply Exercise 3.202 with $K = \Re^n_+$.

3.231

$$K^* = \{\, \mathbf{p} : \mathbf{p}^T\mathbf{x} \leq 0 \text{ for every } \mathbf{x} \in K \,\}$$

No such hyperplane exists if and only if $K^* \cap \Re^n_{++} = \emptyset$. Assume this is the case. By Exercise 3.225, there exists $\mathbf{x} \gneq 0$ such that

$$\mathbf{x}\mathbf{p} = \mathbf{p}^T\mathbf{x} \leq 0 \text{ for every } \mathbf{p} \in K^*$$

In other words, $\mathbf{x} \in K^{**}$. By the duality theorem $K^{**} = K$ which implies that $\mathbf{x} \in K$ as well as $\Re^n_+$, contrary to the hypothesis that $K \cap \Re^n_+ = \{\mathbf{0}\}$. This contradiction establishes that $K^* \cap \Re^n_{++} \neq \emptyset$.

3.232 Given a set of financial assets with prices $\mathbf{p}$ and payoff matrix R, let

$$Z = \{\, (-\mathbf{p}\mathbf{x}, Rx) : \mathbf{x} \in \Re^n \,\}$$

Z is the set of all possible (cost, payoff) pairs. It is a subspace of $\Re^{S+1}$. Let N be the nonnegative orthant in $\Re^{S+1}$. The no arbitrage condition

$$R\mathbf{x} \geq 0 \implies \mathbf{p}^T\mathbf{x} \geq 0$$

implies that $Z \cap N = \{0\}$. By Exercise 3.230, there exists a hyperplane with positive normal $\lambda = \lambda_0, \lambda_1, \dots, \lambda_S$ such that

$$\begin{aligned} \lambda\mathbf{z} &= 0 \qquad \text{for every } \ \mathbf{z} \in Z \\ \lambda\mathbf{z} &> 0 \qquad \text{for every } \ \mathbf{z} \in \Re^{S+1}_+ \setminus \{\mathbf{0}\} \end{aligned}$$

That is

$$-\lambda_0\mathbf{p}\mathbf{x} + \lambda R\mathbf{x} = 0 \qquad \text{for every } \ \mathbf{x} \in \Re^n$$

or

$$\mathbf{p}^T\mathbf{x} = \lambda/\lambda_0 R\mathbf{x} \qquad \text{for every } \ \mathbf{x} \in \Re^n$$

λ/λ_0 is required state price vector.

Conversely, if a state price vector exists

$$p_a = \sum_{a=1}^{S} R_{as}\pi_s$$

then clearly

$$R\mathbf{x} \geq 0 \implies \mathbf{p}^T\mathbf{x} \geq 0$$

No arbitrage portfolios exist.

3.233 Apply the Farkas lemma to the system

$$\begin{aligned} -A\mathbf{x} &\leq 0 \\ -\mathbf{c}^T\mathbf{x} &> 0 \end{aligned}$$

3.234 The inequality system $A^T\mathbf{y} \geq \mathbf{c}$ has a nonnegative solution if and only if the corresponding system of equations

$$A^T\mathbf{y} - \mathbf{z} = \mathbf{c}$$

has a nonnegative solution $\mathbf{y} \in \Re^m_+$, $\mathbf{z} \in \Re^n_+$. This is equivalent to the system

$$B' \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{c} \tag{3.32}$$

where $B' = (A^T, -I_n)$ and I_n is the $n \times n$ identity matrix. By the Farkas lemma, system (3.32) has no solution if and only if the system

$$B\mathbf{x} \leq 0 \text{ and } \ \mathbf{c}^T\mathbf{x} > 0$$

has a solution $\mathbf{x} \in \Re^n$. Since $B = \begin{pmatrix} A \\ -I \end{pmatrix}$, $B\mathbf{x} \leq 0$ implies

$$A\mathbf{x} \leq 0 \text{ and } \ -I\mathbf{x} \leq 0$$

and the latter inequality implies $\mathbf{x} \in \Re^n_+$. Thus we have established that the system $A^T\mathbf{y} \geq \mathbf{c}$ has no nonnegative solution if and only if

$$A\mathbf{x} \leq 0 \text{ and } \ \mathbf{c}^T\mathbf{x} > 0 \text{ for some } \mathbf{x} \in \Re^n_+$$

3.235 Assume system I has a solution, that is there exists $\hat{\mathbf{x}} \in \Re^n_+$ such that

$$A\hat{\mathbf{x}} = \mathbf{0},\ \mathbf{c}\hat{\mathbf{x}} > 0,\ \hat{\mathbf{x}} \geq \mathbf{0}$$

Then $\mathbf{x} = \hat{\mathbf{x}}/\mathbf{c}\hat{\mathbf{x}}$ satisfies the system

$$A\mathbf{x} = \mathbf{0},\ \mathbf{c}\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0} \tag{3.33}$$

which is equivalent to

$$\mathbf{x}'A^T = \mathbf{0},\ \mathbf{x}\mathbf{c} = 1,\ \mathbf{x} \geq \mathbf{0} \tag{3.34}$$

Suppose $\mathbf{y} \in \Re^m$ satisfies

$$A\mathbf{y} \geq \mathbf{c}$$

Multiplying by $\mathbf{x} \geq \mathbf{0}$ gives

$$\mathbf{x}'A^T\mathbf{y} \geq \mathbf{x}\mathbf{c}$$

Substituting (3.34), this implies the contradiction

$$\mathbf{0} \geq \mathbf{1}$$

We conclude that system II cannot have a solution if I has a solution.

Now, assume system I has *no* solution. System I is equivalent to (3.33) which in turn is equivalent to the system

$$\begin{pmatrix} A \\ \mathbf{c} \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

or

$$B\mathbf{x} = \mathbf{b} \tag{3.35}$$

where $B = \begin{pmatrix} -A \\ \mathbf{c} \end{pmatrix}$ is $(m+1) \times n$ and $\mathbf{b} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in \Re^{m+1}$. If (3.35) has no solution, there exists (by the Farkas alternative) some $\mathbf{z} \in \Re^{m+1}$ such that

$$B'\mathbf{z} \leq \mathbf{0} \text{ and } \mathbf{b}\mathbf{z} > 0$$

Decompose $\mathbf{z}$ into $\mathbf{z} = (\mathbf{y}, z)$ with $\mathbf{y} \in \Re^m$ and $z \in \Re$. The second inequality implies that

$$(\mathbf{0}, 1)'(\mathbf{y}, z) = \mathbf{0}\mathbf{y} + z = z > 0$$

Without loss of generality, we can normalize so that $z = 1$ and $\mathbf{z} = (\mathbf{y}, 1)$.

Now $B' = (-A^T, \mathbf{c})$ and so the first inequality implies that

$$(-A^T, \mathbf{c}) \begin{pmatrix} \mathbf{y} \\ 1 \end{pmatrix} = -A^T\mathbf{y} + \mathbf{c} \leq 0$$

or

$$A^T\mathbf{y} \geq \mathbf{c}$$

We conclude that II has a solution.

3.236 For every linear functional g_j, there exists a vector $\mathbf{a} \in \Re^n$ such that

$$g_j(\mathbf{x}) = \mathbf{a}^j \dot{\mathbf{x}}$$

(Proposition 3.11). Let A^T be the matrix whose rows are $\mathbf{a}^j$, that is

$$A = \begin{pmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \dots \\ \mathbf{a}^m \end{pmatrix}$$

Then, the system of inequalities (3.31) is

$$A^T\mathbf{x} \geq \mathbf{c}$$

where $\mathbf{c} = (c_1, c_2, \dots, c_m)$. By the preceding exercise, this system is consistent if and only there is no solution to the system

$$A\lambda = \mathbf{0} \qquad \mathbf{c}\lambda > 0 \qquad \lambda \geq \mathbf{0}$$

Now

$$A\lambda = \mathbf{0} \iff \sum_{j=1}^{m} \lambda_j g_i = \mathbf{0} \qquad i = 1, 2, \dots, m$$

Therefore, the inequalities (3.31) is consistent if an only if

$$\sum_{j=1}^{m} \lambda_j g_i = \mathbf{0} \implies \sum_{j=1}^{m} \lambda_j c_j \leq 0$$

for every set of nonnegative numbers $\lambda_1, \lambda_2, \dots, \lambda_m$.

3.237 Let B be the $2m \times n$ matrix comprising A and $-A$ as follows

$$B = \begin{pmatrix} A \\ -A \end{pmatrix}$$

Then the Fredholm alternative I

$$A\mathbf{x} = 0 \qquad \mathbf{c}^T\mathbf{x} = 1$$

is equivalent to the system

$$B\mathbf{x} \leq 0 \qquad \mathbf{c}\mathbf{x} > 0 \tag{3.36}$$

By the Farkas alternative theorem, either (3.36) has a solution or there exists $\lambda \in R^{2m}_+$ such that

$$B'\lambda = \mathbf{c} \tag{3.37}$$

Decompose λ into two m-vectors

$$\lambda = (\mu, \delta), \ \mu, \delta \in R^m_+$$

so that (3.37) can be rewritten as

$$B'\lambda = A^T\mu - A^T\delta = A^T(\mu - \delta) = \mathbf{c}$$

Define $\mathbf{y} = \mu - \delta \in \Re^m$ We have established that either (3.36) has a solution or there exists a vector $\mathbf{y} \in \Re^m$ such that

$$A^T\mathbf{y} = \mathbf{c}$$

3.238 Let $\mathbf{a}^j$, $j = 1, 2, \ldots, m$ denote the rows of A. Each $\mathbf{a}^i$ defines linear functional $g_j(x) = \mathbf{a}^j x$ on $\Re^n$, and $\mathbf{c}$ defines another linear functional $f(x) = \mathbf{c}^T\mathbf{x}$. Assume that $f(x) = \mathbf{c}^T\mathbf{x} = 0$ for every $\mathbf{x} \in S$ where

$$S = \{\, \mathbf{x} : g_j(\mathbf{x}) = \mathbf{a}^i\mathbf{x} = 0,\ j = 1, 2, \ldots, m \,\}$$

Then the system

$$Ax = 0$$

has no solution satisfying the constraint $\mathbf{c}^T\mathbf{x} > 0$. By Exercise 3.20, there exists scalars $y_1, y_2, \ldots, y_m$ such that

$$f(\mathbf{x}) = \sum_{j=1}^{m} y_j g_j(\mathbf{x})$$

or

$$\mathbf{c} = \sum_{j=1}^{m} y_j a^j = A^T\mathbf{y}$$

That is $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ solves the related nonhomogeneous system

$$A^T\mathbf{y} = \mathbf{c}$$

Conversely, assume that $A^T\mathbf{y} = \mathbf{c}$ for some $y \in \Re^m$. Then

$$\mathbf{c}^T\mathbf{x} = yAx = 0$$

for all x such that $Ax = 0$ and therefore there is no solution satisfying the constraint $\mathbf{c}^T\mathbf{x} = 1$.

3.239 Let

$$S = \{\, \mathbf{z} : \mathbf{z} = A\mathbf{x},\ \mathbf{x} \in \Re \,\}$$

the image of S. S is a subspace. Assume that system I has no solution, that is

$$S \cap \Re^m_{++} = \emptyset$$

By Exercise 3.225, there exists $\mathbf{y} \in \Re^m_+ \setminus \{\mathbf{0}\}$ such that

$$\mathbf{yz} = 0 \text{ for every } \ \mathbf{z} \in S$$

That is

$$\mathbf{y}A\mathbf{x} = 0 \text{ for every } \ \mathbf{x} \in \Re^n$$

Letting $\mathbf{x} = A^T\mathbf{y}$, we have $\mathbf{y}AA^T\mathbf{y} = 0$ which implies that

$$A^T\mathbf{y} = \mathbf{0}$$

System II has a solution $\mathbf{y}$.

Conversely, assume that $\hat{\mathbf{x}}$ is a solution to I. Suppose to the contrary there also exists a solution $\hat{\mathbf{y}}$ to II. Then, since $A\hat{\mathbf{x}} > 0$ and $\hat{\mathbf{y}} \gneq 0$, we must have $\hat{\mathbf{y}}A\hat{\mathbf{x}} = \hat{\mathbf{x}}A^T\hat{\mathbf{y}} > 0$. On the other hand, $A^T\hat{\mathbf{y}} = \mathbf{0}$ which implies $\hat{\mathbf{x}}A^T\hat{\mathbf{y}} = 0$, a contradiction. Hence, we conclude that II cannot have a solution if I has a solution.

3.240 We have already shown (Exercise 3.239) that the alternatives I and II are mutually incompatible. If Gordan's system II

$$A^T\mathbf{y} = \mathbf{0}$$

has a *semipositive* solution $\mathbf{y} \gneq \mathbf{0}$, then we can normalize $\mathbf{y}$ such that $\mathbf{1y} = 1$ and the system

$$\begin{aligned} A^T\mathbf{y} &= \mathbf{0} \\ \mathbf{1y} &= 1 \end{aligned}$$

has a nonnegative solution.

Conversely, if Gordan's system II has *no solution*, the system

$$B'\mathbf{y} = \mathbf{c}$$

where $B' = \begin{pmatrix} A^T \\ \mathbf{1} \end{pmatrix}$ and $\mathbf{c} = (\mathbf{0}, 1) = (0, 0, \dots, 0, 1)$, $\mathbf{0} \in \Re^m$, is the $(m+1)$st unit vector has no solution $\mathbf{y} \geq 0$. By the Farkas lemma, there exists $\mathbf{z} \in \Re^{n+1}$ such that

$$\begin{aligned} B\mathbf{z} &\geq 0 \\ \mathbf{cz} &< 0 \end{aligned}$$

Decompose $\mathbf{z}$ into $\mathbf{z} = (\mathbf{x}, x)$ with $\mathbf{x} \in \Re^n$. The second inequality implies that $x < 0$ since

$$\mathbf{cz} = (\mathbf{0}, 1)'(\mathbf{x}, x) = x < 0$$

Since $B = (A, \mathbf{1})$, the first inequality implies that

$$B\mathbf{z} = (A, \mathbf{1})(\mathbf{x}, x) = A\mathbf{x} + \mathbf{1}x \geq 0$$

or

$$A\mathbf{x} \geq -\mathbf{1}x > 0$$

$\mathbf{x}$ solves Gordan's system I.

3.241 Let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ be a basis for S. Let

$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$$

be the matrix whose columns are $\mathbf{a}_j$. To say that S contains no positive vector means that the system

$$A\mathbf{x} > \mathbf{0}$$

has no solution. By Gordan's theorem, there exists some $\mathbf{y} \gneq \mathbf{0}$ such that

$$A^T\mathbf{y} = \mathbf{0}$$

that is

$$\mathbf{a}_j\mathbf{y} = \mathbf{y}\mathbf{a}_j = 0, \; j = 1, 2, \dots, m$$

so that $\mathbf{y} \in S^\perp$.

3.242 Let Z be the subspace $Z = \{ \mathbf{z} : A\mathbf{x} : \mathbf{x} \in \Re^n \}$. System I has no solution $A\mathbf{x} \gneq \mathbf{0}$ if and only if Z has no nonnegative vector $\mathbf{z} \gneq 0$. By the previous exercise, $Z^\perp$ contains a positive vector $\mathbf{y} > 0$ such that

$$\mathbf{yz} = 0 \text{ for every } \ \mathbf{z} \in Z$$

Letting $\mathbf{x} = A^T\mathbf{y}$, we have $\mathbf{y}AA^T\mathbf{y} = 0$ which implies that

$$A^T\mathbf{y} = \mathbf{0}$$

System II has a solution $\mathbf{y}$.

3.243 Let

$$S = \{ \mathbf{z} : \mathbf{z} = A\mathbf{x}, \ \mathbf{x} \in \Re \}$$

the image of S. S is a subspace. Assume that system I has no solution, that is

$$S \cap \Re^m_+ = \{\mathbf{0}\}$$

By Exercise 3.230, there exists $\mathbf{y} \in \Re^m_{++}$ such that

$$\mathbf{yz} = 0 \text{ for every } \ \mathbf{z} \in S$$

That is

$$\mathbf{y}A\mathbf{x} = 0 \text{ for every } \ \mathbf{x} \in \Re^n$$

Letting $\mathbf{x} = A^T\mathbf{y}$, we have $\mathbf{y}AA^T\mathbf{y} = 0$ which implies that

$$A^T\mathbf{y} = \mathbf{0}$$

System II has a solution $\mathbf{y}$.

Conversely, assume that $\hat{\mathbf{x}}$ is a solution to I. Suppose to the contrary there also exists a solution $\hat{\mathbf{y}}$ to II. Then, since $A\hat{\mathbf{x}} \gneq \mathbf{0}$ and $\hat{\mathbf{y}} > 0$, we must have $\hat{\mathbf{y}}A\hat{\mathbf{x}} = \hat{\mathbf{x}}A^T\hat{\mathbf{y}} > 0$. On the other hand, $A^T\hat{\mathbf{y}} = \mathbf{0}$ which implies $\hat{\mathbf{x}}A^T\hat{\mathbf{y}} = 0$, a contradiction. Hence, we conclude that II cannot have a solution if I has a solution.

3.244 The inequality system $A^T\mathbf{y} \leq \mathbf{0}$ has a nonnegative solution if and only if the corresponding system of equations

$$A^T\mathbf{y} + \mathbf{z} = \mathbf{0}$$

has a nonnegative solution $\mathbf{y} \in \Re^m_+$, $\mathbf{z} \in \Re^n_+$. This is equivalent to the system

$$B' \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{0} \tag{3.38}$$

where $B' = (A^T, I_n)$ and I_n is the $n \times n$ identity matrix. By Gordan's theorem, system (3.38) has no solution if and only if the system

$$B\mathbf{x} > \mathbf{0}$$

has a solution $\mathbf{x} \in \Re^n$. Since $B = \begin{pmatrix} A \\ I \end{pmatrix}$, $B\mathbf{x} > 0$ implies

$$A\mathbf{x} > 0 \text{ and } \ I\mathbf{x} > 0$$

and the latter inequality implies $\mathbf{x} \in \Re^n_{++}$. Thus we have established that the system $A^T\mathbf{y} \leq \mathbf{0}$ has no nonnegative solution if and only if

$$A\mathbf{x} > \mathbf{0} \text{ for some } \mathbf{x} \in \Re^n_{++}$$

3.245 Assume system II has no solution, that is there is no $\mathbf{y} \in \Re^n$ such that

$$A\mathbf{y} \leq 0, \ \mathbf{y} \gneq \mathbf{0}$$

This implies that the system

$$\begin{aligned} -A\mathbf{y} &\geq \mathbf{0} \\ \mathbf{1}\mathbf{y} &\geq 1 \end{aligned}$$

has no solution $\mathbf{y} \in \Re^m_+$. Defining $B' = \begin{pmatrix} -A \\ \mathbf{1}' \end{pmatrix}$, the latter can be written as

$$B'\mathbf{y} \geq -\mathbf{e}_{m+1} \tag{3.39}$$

where $-\mathbf{e}_{m+1} = (\mathbf{0}, 1)$, $\mathbf{0} \in \Re^m$.

By the Gale alternative (Exercise 3.234), if system (3.39) has no solution, the alternative system

$$B\mathbf{z} \leq \mathbf{0}, \ -\mathbf{e}_{m+1}\mathbf{z} > 0$$

has a nonnegative solution $\mathbf{z} \in \Re^{n+1}_+$. Decompose $\mathbf{z}$ into $\mathbf{z} = (\mathbf{x}, z)$ where $\mathbf{x} \in \Re^n_+$ and $z \in \Re_+$. The second inequality implies $z > 0$ since $\mathbf{e}_{m+1}\mathbf{z} = z$.

$B = (-A^T, \mathbf{1})$ and the first inequality implies

$$B\mathbf{z} = (-A^T, \mathbf{1}) \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} = -A^T\mathbf{x} + \mathbf{1}z \leq 0$$

or

$$A^T\mathbf{x} \geq \mathbf{1}z > \mathbf{0}$$

Thus system I has a solution $\mathbf{x} \in \Re^n_+$. Since $\mathbf{x} = \mathbf{0}$ implies $A\mathbf{x} = \mathbf{0}$, we conclude that $\mathbf{x} \gneq \mathbf{0}$.

Conversely, assume that II has a solution $\mathbf{y} \gneq \mathbf{0}$ such that $A\mathbf{y} \leq 0$. Then, for every $\mathbf{x} \in \Re^n_+$

$$\mathbf{x}A^T\mathbf{y} = \mathbf{y}'A^T\mathbf{x} \leq 0$$

Since $\mathbf{y} \gneq \mathbf{0}$, this implies

$$A^T\mathbf{x} \leq 0$$

for every $\mathbf{x} \in \Re^n_+$ which contradicts I.

3.246 We give a constructive proof, by proposing an algorithm which will generate the desired decomposition. Assume that $\mathbf{x}$ satisfies $A\mathbf{x} \gneq \mathbf{0}$. Arrange the rows of A such that the positive elements of $A\mathbf{x}$ are listed first. That is, decompose A into two submatrices such that

$$\begin{aligned} B^1\mathbf{x} &> \mathbf{0} \\ C^1\mathbf{x} &= \mathbf{0} \end{aligned}$$

Either

Case 1 $C^1\mathbf{x} \gneq \mathbf{0}$ has no solution and the result is proved or

Case 2 $C^1\mathbf{x} \gneq \mathbf{0}$ has a solution $\mathbf{x}'$.

Let $\bar{\mathbf{x}}$ be a linear combination of $\mathbf{x}$ and $\mathbf{x}'$. Specifically, define

$$\bar{\mathbf{x}} = \alpha\mathbf{x} + \mathbf{x}'$$

where

$$\alpha > \max \frac{-\mathbf{b}^j\mathbf{x}}{\mathbf{b}^j\mathbf{x}'}$$

where $\mathbf{b}^j$ is the jth row of B^1. α is chosen so that

$$\alpha B^1\mathbf{x} > B^1\mathbf{x}'$$

By direct computation

$$B^1\bar{\mathbf{x}} = \alpha B^1\mathbf{x} + B^1\mathbf{x}' > \mathbf{0}$$
$$C^1\bar{\mathbf{x}} = \alpha C^1\mathbf{x} + C^1\mathbf{x}' \gneq \mathbf{0}$$

since $C^1\mathbf{x} = \mathbf{0}$ and $C^1\mathbf{x}' \gneq \mathbf{0}$. By construction, $\bar{\mathbf{x}}$ is another solution to $A\mathbf{x} \gneq \mathbf{0}$ such that $A\bar{\mathbf{x}}$ has more positive components than $A\mathbf{x}$. Again, collect all the positive components together, decomposing A into two submatrices such that

$$B^2\bar{\mathbf{x}} > \mathbf{0}$$
$$C^2\bar{\mathbf{x}} = \mathbf{0}$$

Either

Case 1 $C^2\mathbf{x} \gneq \mathbf{0}$ has no solution and the result is proved or

Case 2 $C^2\mathbf{x} \gneq \mathbf{0}$ has a solution $\mathbf{x}''$.

In the second case, we can repeat the previous procedure, generating another decomposition B^3, C^3 and so on. At each stage k, the matrix B^k get larger and C^k smaller. The algorithm must terminate before B^k equals A, since we began with the assumption that $A\mathbf{x} > 0$ has no solution.

3.247 There are three possible cases to consider.

Case 1: $\mathbf{y} = \mathbf{0}$ is the only solution of $A^T\mathbf{y} = \mathbf{0}$. Then $A\mathbf{x} > 0$ has a solution $\mathbf{x}'$ by Gordan's theorem and

$$A\mathbf{x}' + \mathbf{0} > \mathbf{0}$$

Case 2: $A^T\mathbf{y} = \mathbf{0}$ has a positive solution $\mathbf{y} > \mathbf{0}$ Then $\mathbf{0}$ is the only solution $A\mathbf{x} \geq \mathbf{0}$ by Stiemke's theorem and

$$A\mathbf{0} + \mathbf{y} > \mathbf{0}$$

Case 3 $A^T\mathbf{y} = \mathbf{0}$ has a solution $\mathbf{y} \gneq \mathbf{0}$ but $\mathbf{y} \not> \mathbf{0}$. By Gordan's theorem $A\mathbf{x} > 0$ has no solution. By the previous exercise, A can be decomposed into two consistent subsystems

$$B\mathbf{x} > \mathbf{0}$$
$$C\mathbf{x} = \mathbf{0}$$

such that $C\mathbf{x} \gneq \mathbf{0}$ has no solution. Assume that B is $k \times n$ and C is $l \times n$ where $l = m - k$. Applying Stiemke's theorem to C, there exists $\mathbf{z} > 0$, $\mathbf{z} \in \Re^l$. Define $\mathbf{y} \in \Re^m_+$ by

$$y_j = \begin{cases} 0 & j = 1, 2, \dots, k \\ y_j = z_{j-k} & j = k+1, k+2, \dots, m \end{cases}$$

Then $\mathbf{x}$, $\mathbf{y}$ is the desired solution since for every j, $j = 1, 2, \dots, m$ either $y_j > 0$ or $(A\mathbf{x})_j = (B\mathbf{x})_j > 0$.

3.248 Consider the dual pair

$$\begin{pmatrix} A \\ I \end{pmatrix} \mathbf{x} \geq 0 \text{ and } (A^T, I) \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{0},\ \mathbf{y} \geq \mathbf{0},\ \mathbf{z} \geq \mathbf{0}$$

By Tucker's theorem, this has a solution $\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*$ such that

$$A\mathbf{x}^* \geq \mathbf{0},\ \mathbf{x}^* \geq \mathbf{0},\ A^T\mathbf{y}^* + \mathbf{z}^* = \mathbf{0},\ \mathbf{y}^* \geq 0,\ \mathbf{z}^* \geq \mathbf{0}$$
$$A\mathbf{x} + \mathbf{y} > \mathbf{0}$$
$$I\mathbf{x}^* + I\mathbf{z} > \mathbf{0}$$

Substituting $\mathbf{z}^* = -A^T\mathbf{y}^*$ implies

$$A^T\mathbf{y} \leq \mathbf{0}$$

and

$$\mathbf{x} - A^T\mathbf{y}^* > 0$$

3.249 Consider the dual pair

$$A\mathbf{x} \geq \mathbf{0} \text{ and } A^T\mathbf{y} = \mathbf{0},\ \mathbf{y} \geq 0$$

where A is an $m \times n$ matrix. By Tucker's theorem, there exists a pair of solutions $\mathbf{x}^* \in \Re^n$ and $\mathbf{y}^* \in \Re^m$ such that

$$A\mathbf{x}^* + \mathbf{y}^* > \mathbf{0} \tag{3.40}$$

Assume that $A\mathbf{x} > \mathbf{0}$ has no solution (Gordan I). Then there exists some j such that $(A\mathbf{x}^*)_j = 0$ and (3.40) implies that $y^*_j > 0$. Therefore $\mathbf{y}^* \gneq \mathbf{0}$ and solves Gordan II.

Conversely, assume that $A^T\mathbf{y} = \mathbf{0}$ has no solution $\mathbf{y} > 0$ (Stiemke II). Then, there exists some j such that $y^*_j = 0$ and (3.40) implies that $(A\mathbf{x}^*)_j > 0$). Therefore $\mathbf{x}^*$ solves $A\mathbf{x} \gneq \mathbf{0}$ (Stiemke I).

3.250 We have already shown that Farkas I and II are mutually inconsistent. Assume that Farkas system I

$$A\mathbf{x} \geq \mathbf{0},\ \mathbf{c}^T\mathbf{x} < 0$$

has no solution. Define the $(m+1) \times n$ matrix $B = \begin{pmatrix} A \\ -\mathbf{c}' \end{pmatrix}$. Our assumption is that the system

$$B\mathbf{x} \geq \mathbf{0}$$

has no solution with $(B\mathbf{x})_{m+1} = -\mathbf{c}\mathbf{x} > 0$. By Tucker's theorem, the dual system

$$B'\mathbf{z} = \mathbf{0}$$

has a solution $\mathbf{z} \in \Re^{m+1}_+$ with $\mathbf{z}_{m+1} > 0$. Without loss of generality, we can normalize so that $\mathbf{z}_{m+1} = 1$. Decompose $\mathbf{z}$ into $\mathbf{z} = (\mathbf{y}, 1)$ with $\mathbf{y} \in \Re^m_+$. Since $B' = (A^T, -\mathbf{c})$, $B'\mathbf{z} = \mathbf{0}$ implies

$$B'\mathbf{z} = (A^T, -\mathbf{c})(\mathbf{y}, 1) = A^T\mathbf{y} - \mathbf{c} = \mathbf{0}$$

or

$$A^T\mathbf{y} = \mathbf{c}$$

$\mathbf{y} \in \Re^m_+$ solves Farkas II.

3.251 If $\mathbf{x} \geq \mathbf{0}$ solves I, then

$$\mathbf{x}'(A^T\mathbf{y}_1 + B'\mathbf{y}_2 + C'\mathbf{y}_3) = \mathbf{x}'A^T\mathbf{y}_1 + \mathbf{x}'B'\mathbf{y}_2 + \mathbf{x}'C'\mathbf{y}_3) > 0$$

since $\mathbf{x}'A^T\mathbf{y}_1 = \mathbf{y}_1 A\mathbf{x} > 0$, $\mathbf{x}'B'\mathbf{y}_2 = \mathbf{y}_2 B\mathbf{x} \geq 0$ and $\mathbf{x}'C'\mathbf{y}_3 = \mathbf{y}_3 C\mathbf{x} = \mathbf{0}$ which contradicts II.

The equation $C\mathbf{x} = \mathbf{0}$ is equivalent to the pair of inequalities $C\mathbf{x} \geq \mathbf{0}, -C\mathbf{x} \geq \mathbf{0}$. By Tucker's theorem the dual pair

$$\begin{aligned} A\mathbf{x} &\geq 0 \qquad A^T\mathbf{y}_1 + B'\mathbf{y}_2 + C'\mathbf{y}_3 - C'\mathbf{y}_4 = \mathbf{0} \\ B\mathbf{x} &\geq \mathbf{0} \\ C\mathbf{x} &\geq \mathbf{0} \\ -C\mathbf{x} &\geq \mathbf{0} \end{aligned}$$

has solutions $x \in \Re^n$, $\mathbf{y}_1 \in \Re^{m_1}$, $\mathbf{y}_2 \in \Re^{m_2}$, $\mathbf{u}_3, \mathbf{v}_3 \in \Re^{m_3}$ such that

$$\begin{aligned} \mathbf{y}_1 &\geq \mathbf{0} & A\mathbf{x} + \mathbf{y}_1 &> \mathbf{0} \\ \mathbf{y}_2 &\geq \mathbf{0} & B\mathbf{x} + \mathbf{y}_2 &> \mathbf{0} \\ \mathbf{u}_3 &\geq \mathbf{0} & C\mathbf{x} + \mathbf{u}_3 &> \mathbf{0} \\ \mathbf{v}_3 &\geq \mathbf{0} & -C\mathbf{x} + \mathbf{v}_3 &> \mathbf{0} \end{aligned}$$

Assume Motzkin I has no solution. That is, there is $\mathbf{y}_1 \gneq \mathbf{0}$. Define $\mathbf{y}_3 = \mathbf{u}_3 - \mathbf{v}_3$. Then $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ satisfies Motzkin II.

3.252 1. For every $\mathbf{a} \in S$, let $S^*_{\mathbf{a}}$ be the polar set

$$S^*_{\mathbf{a}} = \{\, \mathbf{x} \in \Re^n : \|\mathbf{x}\| = 1, \mathbf{x}\mathbf{a} \geq 0 \,\}$$

$S^*_{\mathbf{a}}$ is nonempty since $\mathbf{0} \in S^*_{\mathbf{a}}$. Let $\mathbf{x}$ be the limit of a sequence $\mathbf{x}^n$ of points in $S^*_{\mathbf{a}}$. Since $\mathbf{x}^n\mathbf{a} \geq 0$ for every n, $\mathbf{x}\mathbf{a} \geq 0$ so that $\mathbf{x} \in S^*_{\mathbf{a}}$. Hence $S^*_{\mathbf{a}}$ is a closed subset of $B = \{\, \mathbf{x} \in \Re^n : \|\mathbf{x}\| = 1 \,\}$.

2. Let $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$ be any finite set of points in S. Since $\mathbf{0} \notin S$, the system

$$\sum_{i=1}^m y_i \mathbf{a}_i = \mathbf{0}, \ \sum_{i=1}^m y_i = 1, \ y_i \geq 0$$

has no solution. *A fortiori*, the system

$$\sum_{i=1}^m y_i \mathbf{a}_i = \mathbf{0}$$

has no solution $y \in \Re^m_+$. If A is the $m \times n$ matrix whose rows are $\mathbf{a}_i$, the latter system can be written as

$$A^T \mathbf{y} = \mathbf{0}$$

3. By Gordan's theorem, the system

$$A\mathbf{x} > \mathbf{0} \tag{3.41}$$

has a solution $\bar{\mathbf{x}} \neq \mathbf{0}$.

4. Without loss of generality, we can take $\|\bar{\mathbf{x}}\| = 1$. (3.41) implies that

$$\mathbf{a}_i \bar{\mathbf{x}} = \bar{\mathbf{x}} \mathbf{a}_i > 0$$

for every $i = 1, 2 \dots, m$ so that $\bar{\mathbf{x}} \in S^*_{\mathbf{a}_i}$. Hence

$$\bar{\mathbf{x}} \in \bigcap_{i=1}^m S^*_{\mathbf{a}_i}$$

5. We have shown that for every finite set $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\} \subseteq S$, $\bigcap_{i=1}^m S^*_{\mathbf{a}_i}$ is nonempty closed subset of the compact set $B = \{x \in \Re^n : \|\mathbf{x}\| = 1\}$. By the Finite intersection property (Exercise 1.116)

$$\bigcap_{\mathbf{a} \in S} S^*_{\mathbf{a}} \neq \emptyset$$

6. For every $\mathbf{p} \in \bigcap_{\mathbf{a} \in S} S^*_{\mathbf{a}}$

$$\mathbf{pa} \geq 0 \text{ for every } \mathbf{a} \in S$$

$\mathbf{p}$ defines a hyperplane $f(\mathbf{a}) = \mathbf{pa}$ which separates S from $\mathbf{0}$.

3.253 The expected outcome if player 1 adopts the mixed strategy $\mathbf{p} = (p_1, p_2, \dots, p_m)$ and player 2 plays her j pure strategy is

$$u(\mathbf{p}, j) = \sum_{i=1}^m p_i a_{ij} = \mathbf{pa}^j$$

where $\mathbf{a}^j$ is the jth column of A. The expected payoff to 1 for all possible responses of player 2 is the vector $(\mathbf{p}A)' = A^T\mathbf{p}$. The mixed strategy $\mathbf{p}$ ensures player 1 a nonnegative security level provided $A^T\mathbf{p} \geq \mathbf{0}$.

Similarly, if 2 adopts the mixed strategy $\mathbf{q} = (q_1, q_2, \dots, q_n)$, the expected payoff to 2 if 1 plays his i strategy is $\mathbf{a}^i\mathbf{q}$ where $\mathbf{a}_i$ is the ith row of A. The expected outcome for all the possible responses of player 1 is the vector $A\mathbf{q}$. The mixed strategy $\mathbf{q}$ ensures player 2 a nonpositive security level provided $A\mathbf{q} \leq \mathbf{0}$.

By the von Neumann alternative theorem (Exercise 3.245), at least one of these alternatives must be true. That is, either

Either I $A^T\mathbf{p} > 0$, $\mathbf{p} \gneqq \mathbf{0}$ for some $\mathbf{p} \in \Re^m$

or II $A\mathbf{q} \leq 0$, $\mathbf{q} \gneqq \mathbf{0}$ for some $\mathbf{q} \in \Re^n$

Since $\mathbf{p} \gneqq \mathbf{0}$ and $\mathbf{q} \gneqq \mathbf{0}$, we can normalize so that $\mathbf{p} \in \Delta^{m-1}$ and $\mathbf{q} \in \Delta^{n-1}$. At least one of the players has a strategy which guarantees she cannot lose.

3.254 1. For any $c \in \Re$, define the game

$$\hat{u}(\mathbf{a}^1, \mathbf{a}^2) = u(\mathbf{a}^1, \mathbf{a}^2) - c$$

with

$$\hat{v}_1 = \max_{\mathbf{p}} \min_j \hat{u}(\mathbf{p}, j) = \max_{\mathbf{p}} \min_j u(\mathbf{p}, j) - c = v_1 - c$$
$$\hat{v}_2 = \min_{\mathbf{q}} \max_i \hat{u}(i, \mathbf{q}) = \min_{\mathbf{q}} \max_i u(i, \mathbf{q}) - c = v_2 - c$$

By the previous exercise,

$$\text{Either } \hat{v}_1 \geq 0 \text{ or } \hat{v}_y \leq 0$$

That is

$$\text{Either } v_1 \geq c \text{ or } v_2 \leq c$$

2. Since this applies for arbitrary $c \in \Re$, it implies that while

$$v_1 \leq v_2$$

and there is no c such that

$$v_1 < c < v_2$$

Therefore, we conclude that $v_1 = v_2$ as required.

3.255 1. The mixed strategies $\mathbf{p}$ of player 1 are elements of the simplex Δ^{m-1}, which is compact (Example 1.110). Since $v_1(\mathbf{p}) = \min_{j=1}^n u(\mathbf{p}, j)$ is continuous (Maximum theorem 2.3), $v_1(\mathbf{p})$ achieves its maximum on Δ^{m-1} (Weierstrass theorem 2.2). That is, there exists $\mathbf{p}^* \in \Delta^{m-1}$ such that

$$v_1 = v_1(\mathbf{p}^*) = \max_{\mathbf{p}} v_1(\mathbf{p})$$

Similarly, there exists $\mathbf{q}^* \in \Delta^{n-1}$ such that

$$v_2 = v_2(\mathbf{q}^*) = \min_{\mathbf{q}} v_2(\mathbf{q})$$

2. Let $u(\mathbf{p}, \mathbf{q})$ denote the expected outcome when player 1 adopts mixed strategy $\mathbf{p}$ and player 2 plays $\mathbf{q}$. That is

$$u(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^m \sum_{j=1}^n p_i q_i a_{ij}$$

Then

$$v = u(\mathbf{p}^*, \mathbf{q}^*) = \max_i u(i, \mathbf{q}^*) \geq \sum_i p_i u(i, \mathbf{q}^*) = u(\mathbf{p}, \mathbf{q}^*) \text{ for every } \mathbf{p} \in \Delta^{m-1}$$

Similarly

$$v = u(\mathbf{p}^*, \mathbf{q}^*) = \min_j u(\mathbf{p}^*, j) \leq \sum_j q_j u(\mathbf{p}^*, j) = u(\mathbf{p}^*, \mathbf{q}) \text{ for every } \mathbf{q} \in \Delta^{n-1}$$

$(\mathbf{p}^*, \mathbf{q}^*)$ is a Nash equilibrium.

3.256 By the Minimax theorem, every finite two person zero-sum game has a value. The previous result shows that this is attained at a Nash equilibrium.

3.257 If player 2 adopts the strategy t_1

$$f_{\mathbf{p}}(t_1) = -p_1 + 2p_2 < 0 \text{ if } p_1 > 2p_2$$

If player 2 adopts the strategy t_5

$$f_{\mathbf{p}}(t_5) = p_1 - 2p_2 < 0 \text{ if } p_1 < 2p_2$$

Therefore

$$v_1(\mathbf{p}) = \min_{z \in Z} f_{\mathbf{p}}(\mathbf{z}) \leq \min\{f_{\mathbf{p}}(t_1), f_{\mathbf{p}}(t_5)\} < 0$$

for every $\mathbf{p}$ such that $p_1 \neq p_2$. Since $p_1 + p_2 = 1$, we conclude that

$$v_1(\mathbf{p}) \begin{cases} = 0 & \mathbf{p} = \mathbf{p}^* = (2/3, 1/3) \\ < 0 & \text{otherwise} \end{cases}$$

We conclude that

$$v_1 = \max_{\mathbf{p}} v_1(\mathbf{p}) = 0$$

which is attained at $\mathbf{p}^* = (2/3, 1/3)$.

3.258

1. $$v_2 = \min_{\mathbf{z} \in Z} \max_{i=1}^{m} z_i$$

   Since Z is compact, $v_2 = 0$ implies there exists $\bar{\mathbf{z}} \in Z$ such that

   $$\max_{i=1}^{m} \bar{z}_i = 0$$

   which implies that $\bar{\mathbf{z}} \leq 0$. Consequently $Z \cap \Re^n_- \neq \emptyset$.

2. Assume to the contrary that there exists

   $$\mathbf{z} \in Z \cap \operatorname{int} \Re^n_-$$

   That is, there exists some strategy $\mathbf{q} \in \Delta^{n-1}$ such that $A\mathbf{q} < \mathbf{0}$ and therefore $v_2 < 0$, contrary to the hypothesis.

3. There exists a hyperplane with nonnegative normal separating Z from $\Re^n_-$ (Exercise 3.227). That is, there exists $\mathbf{p}^* \in \Re^n_+$, $\mathbf{p}^* \neq \mathbf{0}$ such that

   $$f_{\mathbf{p}^*}(\mathbf{z}) \geq 0 \text{ for every } \mathbf{z} \in Z$$

   and therefore

   $$v_1(\mathbf{p}^*) = \min_{\mathbf{z} \in Z} f_{\mathbf{p}^*}(\mathbf{z}) \geq 0$$

   Without loss of generality, we can normalize so that $\sum_{i=1}^n p_i^* = 1$ and therefore $\mathbf{p}^* \in \Delta^{m-1}$.

4. Consequently

$$v_1 = \max_{\mathbf{p}} v_1(\mathbf{p}) \geq v_1(\mathbf{p}^*) \geq 0$$

On the other hand, we know that Z contains a point $\bar{\mathbf{z}} \leq 0$. For every $\mathbf{p} \geq \mathbf{0}$

$$f_{\mathbf{p}}(\bar{\mathbf{z}}) \leq 0$$

and therefore

$$v_1(\mathbf{p}) = \min_{\mathbf{z} \in Z} f_{\mathbf{p}}(\mathbf{z}) \leq f_{\mathbf{p}}(\bar{z}) \leq 0$$

so that

$$v_1 = \max_{\mathbf{p}} v_1(\mathbf{p}) \leq 0$$

We conclude that

$$v_1 = 0 = v_2$$

3.259 Consider the game with the same strategies and the payoff function

$$\hat{u}(\mathbf{a}^1, \mathbf{a}^2) = u(\mathbf{a}^1, \mathbf{a}^2) - c$$

The expected value to player 2 is

$$\hat{v}_2 = \min_{\mathbf{q}} \max_{i} \hat{u}(i, \mathbf{q}) = \min_{\mathbf{q}} \max_{i} u(i, \mathbf{q}) - c = v_2 - c = 0$$

By the previous exercise $\hat{v}_1 = \hat{v}_2 = 0$ and

$$v_1 = \max_{\mathbf{p}} \min_{j} u(\mathbf{p}, j) = \max_{\mathbf{q}} \min_{j} \hat{u}(\mathbf{p}, j) + c = \hat{v}_1 + c = c = v_2$$

3.260 Assume that $\mathbf{p}^1$ and $\mathbf{p}^2$ are both optimal strategies for player 1. Then

$$u(\mathbf{p}^1, \mathbf{q}) \geq v \text{ for every } \mathbf{q} \in \Delta^{n-1}$$
$$u(\mathbf{p}^2, \mathbf{q}) \geq v \text{ for every } \mathbf{q} \in \Delta^{n-1}$$

Let $\bar{\mathbf{p}} = \alpha \mathbf{p}^1, \mathbf{p}^2 + (1 - \alpha)$. Since u is bilinear

$$u(\bar{\mathbf{p}}, \mathbf{q}) = \alpha u(\mathbf{p}^1, \mathbf{q}) + (1 - \alpha) u(\mathbf{p}^2, \mathbf{q}) \geq v \text{ for every } \mathbf{q} \in \Delta^{n-1}$$

Consequently, $\bar{\mathbf{p}}$ is also an optimal strategy for player 1.

3.261 f is the payoff function of some 2 person zero-sum game in which the players have $m + 1$ and $n + 1$ strategies respectively. The result follows from the Minimax Theorem.

3.262 1. The possible partitions of $N = \{1, 2, 3\}$ are:

$$\{1\}, \{2\}, \{3\}$$
$$\{i, j\}, \{k\}, \qquad i, j, k = \in N, i \neq j \neq k$$
$$\{1, 2, 3\}$$

In any partition, at most one coalition can have two or more players, and therefore

$$\sum_{k=1}^{K} w(S_k) \leq 1$$

2. Assume $\mathbf{x} = (x_1, x_2, x_3) \in$ core. Then $\mathbf{x}$ must satisfy the following system of inequalities

$$x_1 + x_2 \geq 1 = w(\{1,2\})$$
$$x_1 + x_3 \geq 1 = w(\{1,3\})$$
$$x_2 + x_3 \geq 1 = w(\{2,3\})$$

which can be summed to yield

$$2(x_1 + x_2 + x_3) \geq 3$$

or

$$x_1 + x_2 + x_3 \geq 3/2$$

which implies that $\mathbf{x}$ exceeds the sum available. This contradiction establishes that the core is empty.

Alternatively, observe that the three person majority game is a simple game with no veto players. By Exercise 1.69, its core is empty.

3.263 Assume that the game (N, w) is not cohesive. Then there exists a partition $\{S_1, S_2, \dots, S_K\}$ of N such that

$$w(N) < \sum_{k=1}^{K} w(S_k)$$

Assume $\mathbf{x} \in$ core. Then

$$\sum_{i \in S_k} x_i \geq w(S_k) \qquad k = 1, 2, \dots, K$$

Since $\{S_1, S_2, \dots, S_K\}$ is a partition

$$\sum_{i \in N} x_i = \sum_{k=1}^{K} \sum_{i \in S_k} x_i \geq \sum_{k=1}^{N} w(S_k) > w(N)$$

which contradicts the assumption that $\mathbf{x} \in$ core. This establishes that cohesivity is necessary for the existence of the core.

To show that cohesivity is not sufficient, we observe that the three person majority game is cohesive, but its core is empty.

3.264 The other balanced families of coalitions in a three player game are

1. $\mathcal{B} = \{N\}$ with weights

$$w(S) = \begin{cases} 1 & S = N \\ 0 & \text{otherwise} \end{cases}$$

2. $\mathcal{B} = \{\{1\}, \{2\}, \{3\}\}$ with weights $w(S) = 1$ for every $S \in \mathcal{B}$
3. $\mathcal{B} = \{\{i\}, \{j, k\}\}$, $i, j, k \in \mathcal{B}$, $i \neq j \neq k$ with weights $w(S) = 1$ for every $S \in \mathcal{B}$

3.265 The following table lists some nontrivial balanced families of coalitions for a four player game. Other balanced families can be obtained by permutation of the players.

	Weights
$\{123\}, \{124\}, \{34\}$	$1/2, 1/2, 1/2$
$\{12\}, \{13\}, \{23\}, \{4\}$	$1/2, 1/2, 1/2, 1$
$\{123\}, \{14\}, \{24\}, \{3\}$	$1/2, 1/2, 1/2, 1/2$
$\{123\}, \{14\}, \{24\}, \{34\}$	$2/3, 1/3, 1/3, 1/3$
$\{123\}, \{124\}, \{134\}, \{234\}$	$1/3, 1/3, 1/3, 1/3$

3.266 Both sides of the expression

$$\mathbf{e}_N = \sum_{S \in \mathcal{B}} \lambda_S \mathbf{e}_S$$

are vectors, with each component corresponding to a particular player. For player i, the $i^t h$ component of $\mathbf{e}_N$ is 1 and the $i^t h$ component of $\mathbf{e}_S$ is 1 if $i \in S$ and 0 otherwise. Therefore, for each player i, the preceding expression can be written

$$\sum_{S \in \mathcal{B} | S \ni i} \lambda_S = 1$$

For each coalition S, the share of the coalition S at the allocation $\mathbf{x}$ is

$$g_S(\mathbf{x}) = \sum i \in S x_i = \mathbf{e}_S \dot{\mathbf{x}} \tag{3.42}$$

The condition

$$g_N = \sum_{S \in \mathcal{B}} \lambda_S g_S$$

means that for every $\mathbf{x} \in X$

$$g_N(\mathbf{x}) = \sum_{S \in \mathcal{B}} \lambda_S g_S(\mathbf{x})$$

Substituting (3.42)

$$\mathbf{e}_N \dot{\mathbf{x}} = \sum_{S \in \mathcal{B}} \lambda_S e_S \dot{\mathbf{x}}$$

which is equivalent to the condition

$$\sum_{S \in \mathcal{B}} \lambda_S \mathbf{e}_S = \mathbf{e}_N$$

3.267 By construction, $\mu \geq 0$. If $\mu = 0$,

$$\sum_{S \subseteq N} \lambda_S g_S - \mu g_N = \mathbf{0}$$

implies that $\lambda_S = 0$ for all S and consequently

$$\sum_{S \subseteq N} \lambda_S w(S) - \mu w(N) \leq 0$$

is trivially satisfied. On the other hand, if $\mu > 0$, we can divide both conditions by μ.)

3.268 Let (N, w_1) and (N, w_2) be balanced games. By the Bondareva-Shapley theorem, they have nonempty cores. Let $\mathbf{x}^1 \in \text{core}(N, w_1)$ and $\mathbf{x}^2 \in \text{core}(N, w_2)$. That is,

$$g_S(\mathbf{x}^1) \geq w_1(S) \quad \text{for every} \quad S \subseteq N$$
$$g_S(\mathbf{x}^2) \geq w_2(S) \quad \text{for every} \quad S \subseteq N$$

Adding, we have

$$g_S(\mathbf{x}^1) + g_S(\mathbf{x}^2) = g_S(\mathbf{x}^1 + \mathbf{x}^2) \geq w_1(S) + w_2(S) \quad \text{for every} \quad S \subseteq N$$

which implies that $\mathbf{x}^1 + \mathbf{x}^2$ belongs to $\text{core}(N, w_1 + w_2)$. Therefore $(N, w_1 + w_2)$ is balanced. Similarly, if $\mathbf{x} \in \text{core}(N, w)$, then $\alpha\mathbf{x}$ belongs to $\text{core}(N, \alpha w)$ for every $\alpha \in \Re_+$. That is $(N, \alpha w)$ is balanced for every $\alpha \in \Re_+$.

3.269 1. Assume otherwise. That is assume there exists some $\mathbf{y} \in A \cap B$. Taking the first n components, this implies that

$$\mathbf{e}_N = \sum_{S \subseteq N} \lambda_s \mathbf{e}_S$$

for some $(\lambda_S \geq 0 : S \subseteq N)$. Let $\mathcal{B} = \{S \subset N \mid \lambda_S > 0\}$ be the set of coalitions with strictly positive weights. Then $\mathcal{B}$ is a balanced family of coalitions with weights λ_S (Exercise 3.266).

However, looking at the last coordinate, $\mathbf{y} \in A \cap B$ implies

$$\sum_{S \in \mathcal{B}} \lambda_s w(S) = w(N) + \epsilon > w(N)$$

which contradicts the assumption that the game is balanced. We conclude that A and B are disjoint if the game is balanced.

2. (a) Substituting $\mathbf{y} = (\mathbf{e}_\emptyset, 0)$ in (3.36) gives

$$(\mathbf{z}, z_0)'(\mathbf{0}, 0) = 0 \geq c$$

which implies that $c \leq 0$.

NOTE *We still have to show that $c \geq 0$.*

(b) Substituting $(\mathbf{e}_N, wy(N))$ in (3.36) gives

$$z\mathbf{e}_N + z_0 w(N) > z\mathbf{e}_N + z_0 w(N) + z_0\epsilon$$

for all $\epsilon > 0$, which implies that $z_0 < 0$.

3. Without loss of generality, we can normalize so that $z_0 = -1$. Then the separating hyperplane conditions become

$$(\mathbf{z}, -1)'\mathbf{y} \geq 0 \qquad \text{for every} \quad \mathbf{y} \in A \tag{3.43}$$
$$(\mathbf{z}, -1)'(\mathbf{e}_N, w(N) + \epsilon) < 0 \qquad \text{for every} \quad \epsilon > 0 \tag{3.44}$$

For any $S \subseteq N$, $(\mathbf{e}_S, w(S)) \in A$. Substituting $\mathbf{y} = (\mathbf{e}_S, w(S))$ in (3.43) gives

$$\mathbf{e}_S'\mathbf{z} - w(S) \geq 0$$

that is

$$g_S(\mathbf{z}) = \mathbf{e}_S'\mathbf{z} =\geq w(S)$$

while (3.44) implies

$$g_N(\mathbf{z}) = \mathbf{e}_N'\mathbf{z} > w(N) + \epsilon \qquad \text{for every} \quad \epsilon > 0$$

This establishes that $\mathbf{z}$ belongs to the core. Hence the core is nonempty.

3.270 1. Let $\alpha = w(N) - \sum_{i \in N} w_i > 0$ since (N, w) is essential. For every $S \subseteq N$, define

$$w^0(S) = \frac{1}{\alpha}\left(w(S) - \sum_{i \in S} w_i\right)$$

Then

$$w^0(\{i\}) = 0 \text{ for every } \ i \in N$$
$$w^0(N) = 1$$

w^0 is 0–1 normalized.

2. Let $\mathbf{y} \in \text{core}(N, w^0)$. Then for every $S \subseteq N$

$$\sum_{i \in S} y_i \geq w^0(S) \tag{3.45}$$
$$\sum_{i \in N} y_i = 1 \tag{3.46}$$

Let $\mathbf{w} = (w_1, w_2, \dots, w_n)$ where $w_i = w(\{i\})$. Let $\mathbf{x} = \alpha\mathbf{y} + \mathbf{w}$. Using (3.45) and (3.46)

$$\begin{aligned}
\sum_{i \in S} x_i &= \sum_{i \in S} (\alpha y_i + w_i) \\
&= \alpha \sum_{i \in S} y_i + \sum_{i \in S} w_i \\
&\geq \alpha w^0(S) + \sum_{i \in S} w_i \\
&= \alpha \frac{1}{\alpha}\left(w(S) - \sum_{i \in S} w_i\right) + \sum_{i \in S} w_i \\
&= w(S) \\
\sum_{i \in N} x_i &= \sum_{i \in N} (\alpha y_i + w_i) \\
&= \alpha + \sum_{i \in N} w_i \\
&= w(N)
\end{aligned}$$

Therefore, $\mathbf{x} = \alpha\mathbf{y} + \mathbf{w} \in \text{core}(N, w)$. Similarly, we can show that

$$\mathbf{x} \in \text{core}(N, w) \implies \mathbf{y} = \frac{1}{\alpha}(\mathbf{x} - \mathbf{w}) \in \text{core}(N, w^0)$$

and therefore

$$\text{core}(N, w) = \alpha \text{core}(N, w^0) + \mathbf{w}$$

3. This immediately implies

$$\text{core}(N, w) = \emptyset \iff \text{core}(N, w^0) = \emptyset$$

3.271 (N, w) is 0–1 normalized, that is

$$\begin{aligned} &w(\{i\} = 0 \text{ for every } \; i \in N \\ &w(N) = 1 \end{aligned}$$

Consequently, $\mathbf{x}$ belongs to the core of (N, w) if and only if

$$x_i \geq w_i = 0 \tag{3.47}$$

$$\sum_{i \in N} x_i = w(N) = 1 \tag{3.48}$$

$$\sum_{i \in S} x_i \geq w(S) \text{ for every } \; S \in \mathcal{A} \tag{3.49}$$

(3.46) and (3.47) ensure that $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a mixed strategy for player 1 in the two-person zero-sum game. Using this mixed strategy, the expected payoff to player I for any strategy S of player II is

$$u(\mathbf{x}, S) = \sum_{i \in N} x_i u(i, S) = \sum_{i \in S} x_i \frac{1}{w(S)}$$

(3.48) implies

$$u(\mathbf{x}, S) = \sum_{i \in S} x_i \frac{1}{w(S)} \geq 1 \text{ for every } \; S \in \mathcal{A}$$

That is any $\mathbf{x} \in \text{core}(N, w)$ provides a mixed strategy for player I which ensures a payoff at least 1. That is

$$\text{core}(N, w) \neq \emptyset \implies \delta \geq 1$$

Conversely, if the $\delta < 1$, there is no mixed strategy for player I which satisfies (3.48) and consequently no $\mathbf{x}$ which satisfies (3.46), (3.47) and (3.48). In other words, $\text{core}(N, w) = \emptyset$.

3.272 If δ is the value of G, there exists a mixed strategy which will guarantee that II pays no more than δ. That is, there exists numbers $y_S \geq 0$ for every coalition $S \in \mathcal{A}$ such that

$$\sum_{S \in \mathcal{A}} y_S = 1$$

and

$$\sum_{S \in \mathcal{A}} y_S u(i, S) \leq \delta \qquad \text{for every } \; i \in N$$

that is

$$\sum_{\substack{S \in \mathcal{A} \\ S \ni i}} y_S \frac{1}{w(S)} \leq \delta \qquad \text{for every } \; i \in N$$

or

$$\sum_{\substack{S \in \mathcal{A} \\ S \ni i}} \frac{y_S}{\delta w(S)} \leq 1 \qquad \text{for every } \ i \in N \tag{3.50}$$

For each coalition $S \in \mathcal{A}$ let

$$\lambda_S = \frac{y_S}{\delta w(S)}$$

in (3.50)

$$\sum_{\substack{S \in \mathcal{A} \\ S \ni i}} \lambda_S \leq 1$$

Augment the collection $\mathcal{A}$ with the single-player coalitions to form the collection

$$\mathcal{B} = \mathcal{A} \cup \{ \{i\} : i \in N \}$$

and with weights $\{ \lambda_S : S \in \mathcal{A} \}$ and

$$\lambda_{\{i\}} = 1 - \sum_{S \in \mathcal{A}} \lambda_S$$

Then $\mathcal{B}$ is a balanced collection.

Since the game (N, w) is balanced

$$\begin{aligned} 1 = w(N) &\geq \sum_{S \in \mathcal{B}} \lambda_S w(S) \\ &= \sum_{S \in \mathcal{A}} \lambda_S w(S) \\ &= \sum_{S \in \mathcal{B}} \frac{y_S}{\delta w(S)} w(S) \\ &= \frac{1}{\delta} \sum_{S \in \mathcal{B}} y_S \\ &= \frac{1}{\delta} \end{aligned}$$

that is

$$1 \geq \frac{1}{\delta} \tag{3.51}$$

If I plays the mixed strategy $\bar{\mathbf{x}} = (1/n, 1/n, \ldots, 1/n)$, the payoff is

$$u(\bar{\mathbf{x}}, S) = \sum_{i \in N} \frac{1}{nw(S)} = \frac{1}{w(S)} > 0 \text{ for every } \ S \subseteq \mathcal{A}$$

Therefore $\delta > 0$ and (3.51) implies that

$$\delta \geq 1$$

3.273 Assume $\text{core}(N, w) \neq \emptyset$ and let $x \in \text{core}(N, w)$. Then

$$g_S(\mathbf{x}) \geq w(S) \text{ for every } \quad S \subseteq N \tag{3.52}$$

where $g_S = \sum_{i \in S} x_i$ measures the share coalition S at the allocation $\mathbf{x}$.

Let $\mathcal{B}$ be a balanced family of coalitions with weights λ_S. For every $S \in \mathcal{B}$, (3.52) implies

$$\lambda_S g_S(\mathbf{x}) \geq \lambda_S w(S)$$

Summing over all $S \in \mathcal{B}$

$$\sum_{S \in \mathcal{B}} \lambda_S g_S(\mathbf{x}) \geq \sum_{S \in \mathcal{B}} \lambda_S w(S) \tag{3.53}$$

Evaluating the left hand side of this inequality

$$\begin{aligned}
\sum_{S \in \mathcal{B}} \lambda_S g_S(\mathbf{x}) &= \sum_{S \in \mathcal{B}} \lambda \sum_{i \in S} x_i \\
&= \sum_{i \in N} \sum_{\substack{S \in \mathcal{B} \\ S \ni i}} \lambda x_i \\
&= \sum_{i \in N} x_i \sum_{\substack{S \in \mathcal{B} \\ S \ni i}} \lambda \\
&= \sum_{i \in N} x_i \\
&= w(N)
\end{aligned}$$

Substituting this in (3.53) gives

$$w(N) \geq \sum_{S \in \mathcal{B}} \lambda_S w(S)$$

The game is balanced.