Chapter 5: Optimization

5.1 As stated, this problem has no optimal solution. Revenue $f(x)$ increases without bound as the rate of exploitation x gets smaller and smaller. Given any positive exploitation rate x^0, a smaller rate will increase total revenue. Nonexistence arises from inadequacy in modeling the island leaders' problem. For example, the model ignores any costs of extraction and sale. Realistically, we would expect per-unit costs to decrease with volume (increasing returns to scale) at least over lower outputs. Extraction and transaction costs should make vanishingly small rates of output prohibitively expensive and encourage faster utilization. Secondly, even if the government weights future generations equally with the current generation, it would be rational to value current revenue more highly than future revenue and discount future returns. Discounting is appropriate for two reasons

- Current revenues can be invested to provide a future return. There is an opportunity cost (the interest foregone) to delaying extraction and sale.
- Innovation may create substitutes which reduce the future demand for the fertilizer. If the government is risk averse, it has an incentive to accelerate exploitation, trading-off of lower total return against reduced risk.

5.2 Suppose that $\mathbf{x}^*$ is a local optimum which is not a global optimum. That is, there exists a neighborhood S of $\mathbf{x}^*$ such that

$$f(\mathbf{x}^*, \boldsymbol{\theta}) \geq f(\mathbf{x}, \boldsymbol{\theta}) \text{ for every } \mathbf{x} \in S \cap G(\boldsymbol{\theta})$$

and also another point $\mathbf{x}^{**} \in G(\boldsymbol{\theta})$ such that

$$f(\mathbf{x}^{**}, \boldsymbol{\theta}) > f(\mathbf{x}^*, \boldsymbol{\theta})$$

Since $G(\boldsymbol{\theta})$ is convex, there exists $\alpha \in (0, 1)$ such that

$$\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**} \in S \cap G(\boldsymbol{\theta})$$

By concavity of f

$$f(\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**}, \boldsymbol{\theta}) \geq \alpha f(\mathbf{x}^*, \boldsymbol{\theta}) + (1-\alpha) f(\mathbf{x}^{**}, \boldsymbol{\theta}) > f(\mathbf{x}^*, \boldsymbol{\theta})$$

contradicting the assumption that $\mathbf{x}^*$ is a local optimum.

5.3 Suppose that $\mathbf{x}^*$ is a local optimum which is not a global optimum. That is, there exists a neighborhood S of $\mathbf{x}^*$ such that

$$f(\mathbf{x}^*, \boldsymbol{\theta}) \geq f(\mathbf{x}, \boldsymbol{\theta}) \text{ for every } \mathbf{x} \in S \cap G(\boldsymbol{\theta})$$

and also another point $\mathbf{x}^{**} \in G(\boldsymbol{\theta})$ such that

$$f(\mathbf{x}^{**}, \boldsymbol{\theta}) > f(\mathbf{x}^*, \boldsymbol{\theta})$$

Since $G(\boldsymbol{\theta})$ is convex, there exists $\alpha \in (0, 1)$ such that

$$\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**} \in S \cap G(\boldsymbol{\theta})$$

By strict quasiconcavity of f

$$f(\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**}, \boldsymbol{\theta}) > \min\{\, f(\mathbf{x}^*, \boldsymbol{\theta}), f(\mathbf{x}^{**}, \boldsymbol{\theta})\,\} > f(\mathbf{x}^*, \boldsymbol{\theta})$$

contradicting the assumption that $\mathbf{x}^*$ is a local optimum. Therefore, if $\mathbf{x}^*$ is local optimum, it must be a global optimum.

Now suppose that $\mathbf{x}^*$ is a weak global optimum, that is

$$f(\mathbf{x}^*, \boldsymbol{\theta}) \geq f(\mathbf{x}, \boldsymbol{\theta}) \text{ for every } \ \mathbf{x} \in S$$

but there another point $\mathbf{x}^{**} \in S$ such that

$$f(\mathbf{x}^{**}, \boldsymbol{\theta}) = f(\mathbf{x}^*, \boldsymbol{\theta})$$

Since $G(\boldsymbol{\theta})$ is convex, there exists $\alpha \in (0,1)$ such that

$$\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**} \in S \cap G(\boldsymbol{\theta})$$

By strict quasiconcavity of f

$$f(\alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^{**}, \boldsymbol{\theta}) > \min\{\, f(\mathbf{x}^*, \boldsymbol{\theta}), f(\mathbf{x}^{**}, \boldsymbol{\theta})\,\} = f(\mathbf{x}^*, \boldsymbol{\theta})$$

contradicting the assumption that $\mathbf{x}^*$ is a global optimum. We conclude that every optimum is a strict global optimum and hence unique.

5.4 Suppose that $\mathbf{x}^*$ is a local optimum of (5.3) in X, so that

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \tag{5.1}$$

for every $\mathbf{x}$ in a neighborhood S of $\mathbf{x}^*$. If f is differentiable,

$$f(\mathbf{x}) = f(\mathbf{x}^*) + Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) + \eta(\mathbf{x})\,\|\mathbf{x} - \mathbf{x}^*\|$$

where $\eta(\mathbf{x}) \to 0$ as $\mathbf{x} \to \mathbf{x}^*$. (5.1) implies that there exists a ball $B_r(\mathbf{x}^*)$ such that

$$Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) + \eta(\mathbf{x})\,\|\mathbf{x} - \mathbf{x}^*\| \leq 0$$

for every $\mathbf{x} \in B_r(\mathbf{x}^*)$. Letting $\mathbf{x} \to \mathbf{x}^*$, we conclude that

$$Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) \leq 0$$

for every $\mathbf{x} \in B_r(\mathbf{x}^*)$.

Suppose there exists $\mathbf{x} \in B_r(\mathbf{x}^*)$ such that

$$Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) = y < 0$$

Let $\mathbf{dx} = \mathbf{x} - \mathbf{x}^*$ so that $\mathbf{x} = \mathbf{x}^* + \mathbf{dx}$. Then $\mathbf{x}^* - \mathbf{dx} \in B_r(\mathbf{x}^*)$. Since $Df[\mathbf{x}^*]$ is linear,

$$Df[\mathbf{x}^*](-\mathbf{dx}) = -Df[\mathbf{x}^*](\mathbf{dx}) = -y > 0$$

contradicting (5.1). Therefore

$$Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) = 0$$

for every $\mathbf{x} \in B_r(\mathbf{x}^*)$.

5.5 We apply the reasoning of Example 5.5 to each component. Formally, for each i, let $\hat{f}_i$ be the projection of f along the i^{th} axis

$$\hat{f}_i(x_i) = f(x_1^*, x_2^*, \dots, x_{i-1}^*, x_i, x_{i+1}^*, \dots, x_n^*)$$

x_i^* maximizes $\hat{f}_i(x_i)$ over $\Re_+$, for which it is necessary that

$$D_{x_i}\hat{f}_i(x_i^*) \leq 0 \qquad x_i^* \geq 0 \qquad x_i^* D_{x_i}\hat{f}_i(x_i^*) = 0$$

Substituting

$$D_{x_i}\hat{f}_i(x_i^*) = D_{x_i} f[\mathbf{x}^*]$$

yields

$$D_{x_i} f[\mathbf{x}^*] \leq 0 \qquad x_i^* \geq 0 \qquad x_i^* D_{x_i} f[\mathbf{x}^*] = 0$$

5.6 By Taylor's Theorem (Example 4.33)

$$f(\mathbf{x}^* + \mathbf{dx}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)\mathbf{dx} + \frac{1}{2}\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} + \eta(\mathbf{dx})\left\|\mathbf{dx}\right\|^2$$

with $\eta(\mathbf{dx}) \to \mathbf{0}$ as $\mathbf{dx} \to \mathbf{0}$. Given

1. $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and
2. $H_f(\mathbf{x}^*)$ is negative definite

and letting $\mathbf{dx} \to \mathbf{0}$, we conclude that

$$f(\mathbf{x}^* + \mathbf{dx}) < f(\mathbf{x}^*)$$

for small $\mathbf{dx}$. $\mathbf{x}^*$ is a strict local maximum.

5.7 If $\mathbf{x}^*$ is a local minimum of $f(\mathbf{x})$, it is necessary that

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$

for every $\mathbf{x}$ in a neighborhood S of $\mathbf{x}^*$. Assuming that f is C^2, $f(\mathbf{x})$ can be approximated by

$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)\mathbf{dx} + \frac{1}{2}\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx}$$

where $\mathbf{dx} = \mathbf{x} - \mathbf{x}^*$. If $\mathbf{x}^*$ is a local minimum, then there exists a ball $B_r(\mathbf{x}^*)$ such that

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)\mathbf{dx} + \frac{1}{2}\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx}$$

or

$$\nabla f(\mathbf{x}^*)\mathbf{dx} + \frac{1}{2}\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} \geq 0$$

for every $\mathbf{dx} \in B_r(\mathbf{x}^*)$. To satisfy this inequality for all small $\mathbf{dx}$ requires that the first term be zero and the second term nonnegative. In other words, for a point $\mathbf{x}^*$ to be a local minimum of a function f, it is necessary that the gradient be zero and the Hessian be nonnegative definite at $\mathbf{x}^*$. Furthermore, by Taylor's Theorem

$$f(\mathbf{x}^* + \mathbf{dx}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)\mathbf{dx} + \frac{1}{2}\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} + \eta(\mathbf{dx})\left\|\mathbf{dx}\right\|^2$$

with $\eta(\mathbf{dx}) \to \mathbf{0}$ as $\mathbf{dx} \to \mathbf{0}$. Given

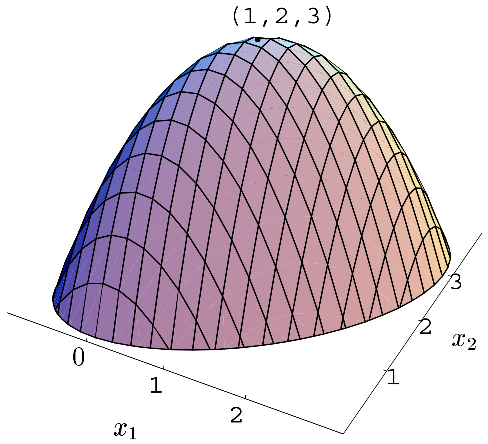

Figure 5.1: The strictly concave function $f(x_1, x_2) = x_1x_2 + 3x_2 - x_1^2 - x_2^2$ has a unique global maximum.

1. $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and
2. $H_f(\mathbf{x}^*)$ is positive definite

and letting $\mathbf{dx} \to \mathbf{0}$, we conclude that

$$f(\mathbf{x}^* + \mathbf{dx}) > f(\mathbf{x}^*)$$

for small $\mathbf{dx}$. $\mathbf{x}^*$ is a strict local minimum.

5.8 By the Weierstrass theorem (Theorem 2.2), f has a maximum x^* and a minimum x_* on $[a, b]$. Either

- $x^* \in (a, b)$ and $f'(x^*) = 0$ (Theorem 5.1) or
- $x_* \in (a, b)$ and $f'(x^*) = 0$ (Exercise 5.7) or
- Both maxima and minima are boundary points, that is $x^*, x_* \in \{a, b\}$ which implies that f is constant on $[a, b]$ and therefore $f'(x) = 0$ for every $x \in (a, b)$ (Exercise 4.7).

5.9 The first-order conditions for a maximum are

$$D_{x_1} f(x_1, x_2) = x_2 - 2x_1 = 0$$
$$D_{x_2} f(x_1, x_2) = x_1 + 3 - 2x_2 = 0$$

which have the unique solution $x_1^* = 1, x_2^* = 2$. $(1, 2)$ is the only stationary point of f and hence the only possible candidate for a maximum. To verify that $(1, 2)$ satisfies the second-order condition for a maximum, we compute the Hessian of f

$$H(\mathbf{x}) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

which is negative definite everywhere. Therefore $(1, 2)$ is a strict local maximum of f. Further, since f is strictly concave (Proposition 4.1), we conclude that $(1, 2)$ is a strict global maximum of f (Exercise 5.2), where it attains its maximum value $f(1, 2) = 3$ (Figure 5.1).

5.10 The first-order conditions for a maximum (or minimum) are

$$D_1 f(x) = 2x_1 = 0$$
$$D_2 f(x) = 2x_2 = 0$$

which have a unique solution $x_1 = x_2 = 0$. This is the only stationary point of f. Since the Hessian of f

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

is positive definite, we deduce $(0,0)$ is a strict global minimum of f (Proposition 4.1, Exercise 5.2).

5.11 The average firm's profit function is

$$\Pi(k,l) = y - \frac{1}{2}k - l - \frac{1}{6}$$

and the firm's profit maximization problem is

$$\max_{k,l} \Pi(k,l) = k^{1/6}l^{1/3} - \frac{1}{2}k - l - \frac{1}{6}$$

A necessary condition for a profit maximum is that the profit function be stationary, that is

$$D_k\Pi(k,l) = \frac{1}{6}k^{-5/6}l^{1/3} - \frac{1}{2} = 0$$
$$D_l\Pi(k,l) = \frac{1}{3}k^{1/6}l^{-2/3} - 1 = 0$$

which can be solved to yield

$$k = l = \frac{1}{9}$$

The firm's output is

$$y = \frac{1}{9}^{1/6}\frac{1}{9}^{1/3} = \frac{1}{3}$$

and its profit is

$$\Pi(\frac{1}{3},\frac{1}{3}) = \frac{1}{3} - \frac{1}{2}\frac{1}{9} - \frac{1}{9} - \frac{1}{6} = 0$$

5.12 By the Chain Rule

$$D_{\mathbf{x}}(h \circ f)[\mathbf{x}^*] = Dh \circ D_{\mathbf{x}} f[\mathbf{x}^*] = 0$$

Since $Dh > 0$

$$D_{\mathbf{x}}(h \circ f)[x^*] = 0 \iff D_{\mathbf{x}} f[\mathbf{x}^*] = 0$$

$h \circ f$ has the same stationary points as f.

5.13 Since the log function is monotonic, finding the maximum likelihood estimators is equivalent to solving the maximization problem (Exercise 5.12)

$$\max_{\mu,\sigma} \log L(\mu,\sigma) = -\frac{T}{2}\log 2\pi - T\log\sigma - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(x_t-\mu)^2$$

For $(\hat{\mu},\hat{\sigma}^2)$ to solve this problem, it is necessary that $\log L$ be stationary at $(\hat{\mu},\hat{\sigma}^2)$, that is

$$D_\mu \log L(\hat{\mu},\hat{\sigma}^2) = \frac{1}{\hat{\sigma}^2}\sum_{t=1}^{T}(x_t-\hat{\mu}) = 0$$
$$D_\sigma \log L(\hat{\mu},\hat{\sigma}^2) = -\frac{T}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{t=1}^{T}(x_t-\hat{\mu})^2 = 0$$

which can be solved to yield

$$\hat{\mu} = \bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t$$
$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})^2$$

5.14 The gradient of the objective function is

$$\nabla f(\mathbf{x}) = \begin{pmatrix} -2(x_1-1) \\ -2(x_2-1) \end{pmatrix}$$

while that of the constraint is

$$\nabla g(\mathbf{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

A necessary condition for the optimal solution is that these be proportional that is

$$\nabla f(x) = \begin{pmatrix} -2(x_1-1) \\ -2(x_2-1) \end{pmatrix} = \lambda \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \nabla g(\mathbf{x})$$

which can be solved to yield

$$x_1 = x_2 = \frac{1}{1+\lambda}$$

which includes an unknown constant of proportionality λ. However, any solution must also satisfy the constraint

$$g(x_1,x_2) = 2\left(\frac{1}{1+\lambda}\right)^2 = 1$$

This can be solved for λ

$$\lambda = \sqrt{2} - 1$$

and substituted into (5.1)

$$x_1 = x_2 = \frac{1}{\sqrt{2}}$$

5.15 The consumer's problem is

$$\max_{\mathbf{x} \geq 0} u(\mathbf{x}) = x_1 + a \log x_2$$
$$\text{subject to } g(\mathbf{x}) = x_1 + p_2 x_2 - m = 0$$

The first-order conditions for a (local) optimum are

$$D_{x_1} u(\mathbf{x}^*) = 1 \leq \lambda = D_{x_1} g(x^*) \qquad x_1 \geq 0 \qquad x_1(1-\lambda) = 0 \tag{5.2}$$
$$D_{x_2} u(\mathbf{x}^*) = \frac{a}{x_2} \leq \lambda p_2 = D_{x_2} g(x^*) \qquad x_2 \geq 0 \qquad x_2\left(\frac{a}{x_2} - \lambda p_2\right) = 0 \tag{5.3}$$

We can distinguish two cases:

Case 1 $x_1 = 0$ in which case the budget constraint implies that $x_2 = m/p_2$.

Case 2 $x_1 > 0$ In this case, (5.2) implies that $\lambda = 1$. Consequently, the first inequality of (5.3) implies that $x_2 > 0$ and therefore the last equation implies $x_2 = a/p_2$ with $x_1 = m - a$.

We deduce that the consumer first spends portion a of her income on good 2 and the remainder on good 1.

5.16 Suppose without loss of generality that the first k components of $\mathbf{y}^*$ are strictly positive while the remaining components are zero. That is

$$\begin{aligned} y_i^* &> 0 \qquad i = 1, 2, \ldots, k \\ y_i^* &= 0 \qquad i = k+1, k+2, \ldots, n \end{aligned}$$

$(\mathbf{x}^*, \mathbf{y}^*)$ solves the problem

$$\max f(\mathbf{x})$$
$$\text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}$$
$$y_i = 0 \qquad i = k+1, k+2, \ldots, n$$

By Theorem 5.2, there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and $\mu_{k+1}, \mu_{k+2}, \ldots, \mu_n$ such that

$$D_{\mathbf{x}} f[\mathbf{x}^*, \mathbf{y}^*] = \sum_{j=1}^{m} \lambda_j D_{\mathbf{x}} g_j[\mathbf{x}^*, \mathbf{y}^*]$$
$$D_{\mathbf{y}} f[\mathbf{x}^*, \mathbf{y}^*] = \sum_{j=1}^{m} \lambda_j D_{\mathbf{y}} g_j[\mathbf{x}^*, \mathbf{y}^*] + \sum_{i=k+1}^{n} \mu_i y_i$$

Furthermore, $\mu_i \geq 0$ for every i so that

$$D_{\mathbf{y}} f[\mathbf{x}^*, \mathbf{y}^*] \leq \sum_{j=1}^{m} \lambda_j D_{\mathbf{y}} g_j[\mathbf{x}^*, \mathbf{y}^*]$$

with

$$D_{y_i} f[\mathbf{x}^*, \mathbf{y}^*] = \sum_{j=1}^{m} \lambda_j D_{y_i} g_j[\mathbf{x}^*, \mathbf{y}^*] \text{ if } y_i > 0$$

5.17 Assume that $\mathbf{x}^* = (x_1^*, x_2^*)$ solves

$$\max_{x_1, x_2} f(x_1, x_2)$$

subject to

$$g(x_1, x_2) = 0$$

By the implicit function theorem, there exists a function $h : \Re \to \Re$ such that

$$x_1 = h(x_2) \tag{5.4}$$

and

$$g(h(x_2), x_2) = 0$$

for x_2 in a neighborhood of x_2^*. Furthermore

$$Dh[x_2^*] = -\frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]} \tag{5.5}$$

Using (5.4), we can convert the original problem into the unconstrained maximization of a function of a single variable

$$\max_{x_2} f(h(x_2), x_2)$$

If x_2^* maximizes this function, it must satisfy the first-order condition (applying the Chain Rule)

$$D_{x_1} f[x*] \circ Dh[x_2^*] + D_{x_2} f[\mathbf{x}^*] = 0$$

Substituting (5.5) yields

$$D_{x_1} f[x*] \left(-\frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]} \right) + D_{x_2} f[\mathbf{x}^*] = 0$$

or

$$\frac{D_{x_1} f[x*]}{D_{x_2} f[\mathbf{x}^*]} = \frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]}$$

5.18 The consumer's problem is

$$\max_{\mathbf{x} \in X} u(\mathbf{x})$$
$$\text{subject to } \mathbf{p}^T \mathbf{x} = m$$

Solving for x_1 from the budget constraint yields

$$x_1 = \frac{m - \sum_{i=2}^n p_i x_i}{p_1}$$

Substituting this in the utility function, the affordable utility levels are

$$\hat{u}(x_2, x_3, \ldots, x_n) = u\left(\frac{m - \sum_{i=2}^n p_i x_i}{p_1}, x_2, x_3, \ldots, x_n \right) \tag{5.6}$$

and the consumer's problem is to choose $(x_2, x_3, \ldots, x_n)$ to maximize (5.6). The first-order conditions are that $\hat{u}(x_2, x_3, \ldots, x_n)$ be stationary, that is for every good $j = 2, 3, \ldots, n$

$$D_{x_j} \hat{u}(x_2, x_3, \ldots, x_n) = D_{x_1} u(\mathbf{x}^*) D_{x_j} \left(\frac{m - \sum_{i=2}^n p_i x_i}{p_1} \right) + D_{x_j} u(\mathbf{x}^*) = 0$$

which reduces to

$$D_{x_1}u(\mathbf{x}^*)(-\frac{p_1}{p_j}) + D_{x_j}u(\mathbf{x}^*) = 0$$

or

$$\frac{D_{x_1}u(\mathbf{x}^*)}{D_{x_j}u(\mathbf{x}^*)} = \frac{p_1}{p_j} \qquad j = 2, 3, \dots, n$$

This is the familiar equality between the marginal rate of substitution and the price ratio (Example 5.15). Since our selection of x_1 was arbitrary, this applies between any two goods.

5.19 Adapt Exercise 5.6.

5.20 Corollary 5.1.2 implies that $\mathbf{x}^*$ is a global maximum of $L(\mathbf{x}, \boldsymbol{\lambda})$, that is

$$L(\mathbf{x}^*, \boldsymbol{\lambda}) \geq L(\mathbf{x}, \boldsymbol{\lambda}) \text{ for every } \ \mathbf{x} \in X$$

which implies

$$f(\mathbf{x}^*) - \sum \lambda_j g_j(\mathbf{x}^*) \geq f(\mathbf{x}) - \sum \lambda_j g_j(\mathbf{x}) \text{ for every } \ \mathbf{x} \in X$$

Since $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ this implies

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) - \sum \lambda_j g_j(\mathbf{x}) \text{ for every } \ \mathbf{x} \in X$$

A fortiori

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for every } \ \mathbf{x} \in G = \{\, \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) = \mathbf{0} \,\}$$

5.21 Suppose that $\mathbf{x}^*$ is a local maximum of f on G. That is, there exists a neighborhood S such that

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for every } \ \mathbf{x} \in S \cap G$$

But for every $\mathbf{x} \in G$, $g_j(\mathbf{x}) = 0$ for every j and

$$L(\mathbf{x}) = f(\mathbf{x}) + \sum \lambda_j g_j(\mathbf{x}) = f(\mathbf{x})$$

and therefore

$$L(\mathbf{x}^*) \geq L(\mathbf{x}) \text{ for every } \ \mathbf{x} \in S \cap G$$

5.22 The area of the base is

$$\text{Base} = w^2 = A/3$$

and the four sides

$$\begin{aligned} \text{Sides} &= 4wh \\ &= 4\sqrt{\frac{A}{3}}\sqrt{\frac{A}{12}} \\ &= \frac{4A}{16} \\ &= \frac{2A}{3} \end{aligned}$$

5.23 Let the dimensions of the vat be $w \times l \times h$. We wish to

$$\min_{w,l,h} \text{Surface area} = A = w \times l + 2wh + 2lh$$

$$\text{subject to } w \times l \times h = 32$$

The Lagrangean is

$$L(w, l, h, \lambda) = wl + 2wh + 2lh - \lambda wlh.$$

The first-order conditions for a maximum are

$$\begin{aligned} D_w L &= l + 2h - \lambda lh = 0 && (5.7)\\ D_l L &= w + 2h - \lambda wh = 0 && (5.8)\\ D_h L &= 2w + 2l - \lambda wl = 0 && (5.9)\\ wlh &= 32 \end{aligned}$$

Subtracting (5.8) from (5.7)

$$l - w = \lambda(l - w)h$$

This equation has two possible solutions. Either

$$\lambda = \frac{1}{h} \text{ or } l = w$$

But if $\lambda = 1/h$, (5.7) implies that $l = 0$ and the volume is zero. Therefore, we conclude that $w = l$. Substituting $w = l$ in (5.8) and (5.9) gives

$$\begin{aligned} w + 2h &= \lambda wh\\ 4w &= \lambda w^2 \end{aligned}$$

from which we deduce that

$$\lambda = \frac{4}{w}$$

Substituting in (5.8)

$$w + 2h = \frac{4}{w} wh = 4h$$

which implies that

$$w = 2h \text{ or } h = \frac{1}{2} w$$

To achieve the required volume of 32 cubic metres requires that

$$w \times l \times h = w \times w \times \frac{1}{2} w = 32$$

so that the dimensions of the vat are

$$w = 4 \qquad l = 4 \qquad h = 2$$

The area of sheet metal required is

$$A = wl + 2wh + 2lh = 48$$

5.24 The Lagrangean for this problem is

$$L(\mathbf{x}, \lambda) = x_1^2 + x_2^2 + x_3^2 - \lambda(2x_1 - 3x_2 + 5x_3 - 19)$$

A necessary condition for $\mathbf{x}^*$ to solve the problem is that the Lagrangean be stationary at $\mathbf{x}^*$, that is

$$\begin{aligned} D_{x_1}L &= 2x_1^* - 2\lambda = 0 \\ D_{x_2}L &= 2x_2^* + 3\lambda = 0 \\ D_{x_3}L &= 2x_3^* - 5\lambda = 0 \end{aligned}$$

which implies

$$x_1^* = \lambda \qquad x_2^* = -\frac{3}{2}\lambda \qquad x_2^* = \frac{5}{2}\lambda \tag{5.10}$$

It is also necessary that the solution satisfy the constraint, that is

$$2x_1^* - 3x_2^* + 5x_3^* = 19$$

Substituting (5.10) into the constraint we get

$$2\lambda + \frac{9}{2}\lambda + \frac{25}{2}\lambda = 19\lambda = 19$$

which implies $\lambda = 1$. Substituting in (5.10), the solution is $\mathbf{x}^* = (1, -\frac{3}{2}, \frac{5}{2})$. Since the constraint is affine and the objective $(-f)$ is concave, stationarity of the Lagrangean is also sufficient for global optimum (Corollary 5.2.4).

5.25 The Lagrangean is

$$L(x_1, x_2, \lambda) = x_1^\alpha x_2^{1-\alpha} - \lambda(p_1x_1 + p_2x_2 - m)$$

The Lagrangean is stationary where

$$\begin{aligned} D_{x_1}L &= \alpha x_1^{\alpha-1}x_2^{1-\alpha} - \lambda p_1 = 0 \\ D_{x_2}L &= 1 - \alpha x_1^\alpha x_2^{1-\alpha-1} - \lambda p_2 = 0 \end{aligned}$$

Therefore the first-order conditions for a maximum are

$$\alpha x_1^{\alpha-1}x_2^{1-\alpha} = \lambda p_1 \tag{5.11}$$

$$1 - \alpha x_1^\alpha x_2^{1-\alpha-1} = \lambda p_2 \tag{5.12}$$

$$p_1x_1 + p_2x_2 - m \tag{5.13}$$

Dividing (5.11) by (5.12) gives

$$\frac{\alpha x_1^{\alpha-1}x_2^{1-\alpha}}{1 - \alpha x_1^\alpha x_2^{(1-\alpha)-1}} = p_1p_2$$

which simplifies to

$$\frac{\alpha x_2}{(1-\alpha)x_1} = \frac{p_1}{p_2} \qquad \text{or} \qquad p_2x_2 = \frac{(1-\alpha)}{\alpha}p_1x_1$$

Substituting in the budget constraint (5.13)

$$\begin{aligned} p_1x_1 + \frac{(1-\alpha)}{\alpha}p_1x_1 &= m \\ \frac{\alpha + (1-\alpha)}{\alpha}p_1x_1 &= m \end{aligned}$$

so that

$$x_1^* = \frac{\alpha}{\alpha + (1-\alpha)} \frac{m}{p_1}$$

From the budget constraint (5.13)

$$x_2^* = \frac{(1-\alpha)}{\alpha + (1-\alpha)} \frac{m}{p_2}$$

5.26 The Lagrangean is

$$L(\mathbf{x}, \lambda) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} - \lambda(p_1 x_1 + p_2 x_2 + \dots + p_n x_n)$$

The first-order conditions for a maximum are

$$D_{x_i} L = \alpha_i x_1^{\alpha_1} x_2^{\alpha_2} \dots x_i^{\alpha_i - 1} \dots x_n^{\alpha_n} - \lambda p_i = \frac{\alpha_i u(x)}{x_i} - \lambda p_i = 0$$

or

$$\frac{\alpha_i u(x)}{\lambda} = p_i x_i \qquad i = 1, 2, \dots, n \tag{5.14}$$

Summing over all goods and using the budget constraint

$$\sum_{i=1}^n \frac{\alpha_i u(x)}{\lambda} = \frac{u(x)}{\lambda} \sum_{i=1}^n \alpha_i = \sum_{i=1}^n p_i x_i = m$$

Letting $\sum_{i=1}^n \alpha_i = \alpha$, this implies

$$\frac{u(\mathbf{x})}{\lambda} = \frac{m}{\alpha}$$

Substituting in (5.14)

$$p_i x_i = \frac{\alpha_i}{\alpha} m$$

or

$$x_i^* = \frac{\alpha_i}{\alpha} \frac{m}{p_i} \qquad i = 1, 2, \dots, n$$

5.27 The Lagrangean is

$$L(\mathbf{x}, \lambda) = w_1 x_1 + w_2 x_2 - \lambda(x_1^\rho + x_2^\rho - y^\rho).$$

The necessary conditions for stationarity are

$$\begin{aligned} D_{x_1} L(\mathbf{x}, \lambda) &= w_1 - \lambda \rho x_1^{\rho-1} = 0 \\ D_{x_2} L(\mathbf{x}, \lambda) &= w_2 - \lambda \rho x_2^{\rho-1} = 0 \end{aligned}$$

or

$$\begin{aligned} w_1 &= \lambda \rho x_1^{\rho-1} \\ w_2 &= \lambda \rho x_2^{\rho-1} \end{aligned}$$

which reduce to

$$\frac{w_1}{w_2} = \frac{x_1^{\rho-1}}{x^{\rho-1}{}_2}$$
$$x_2^{\rho-1} = \frac{w_2}{w_1} x_1^{\rho-1}$$
$$x_2^{\rho} = \left(\frac{w_2}{w_1}\right)^{\frac{\rho}{\rho-1}} x_1^{\rho}$$

Substituting in the production constraint

$$x_1^{\rho} + \left(\frac{w_2}{w_1}\right)^{\frac{\rho}{\rho-1}} x_1^{\rho} = y^{\rho}$$
$$\left(1 + \left(\frac{w_2}{w_1}\right)^{\frac{\rho}{\rho-1}}\right) x_1^{\rho} = y^{\rho}$$

we can solve x_1

$$x_1^{\rho} = \left(1 + \left(\frac{w_2}{w_1}\right)^{\frac{\rho}{\rho-1}}\right)^{-1} y^{\rho}$$
$$x_1 = \left(1 + \left(\frac{w_2}{w_1}\right)^{\frac{\rho}{\rho-1}}\right)^{-1/p} y$$

Similarly

$$x_2 = \left(1 + \left(\frac{w_1}{w_2}\right)^{\frac{\rho}{\rho-1}}\right)^{-1/p} y$$

5.28 Example 5.27 is flawed. The optimum of the constrained maximization problem ($h = w/2$) is in fact a saddle point of the Lagrangean. It maximizes the Lagrangean in the feasible set, but not globally.

The net benefit approach to the Lagrange multiplier method is really only applicable when the Lagrangean (net benefit function) is concave, so that every stationary point is a global maximum. This requirement is satisfied in many standard examples, such as the consumer's problem (Example 5.21) and cost minimization (Example 5.28). It is also met in Example 5.29. The requirement of concavity is not recognized in the text, and Section 5.3.6 should be amended accordingly.

5.29 The Lagrangean

$$L(\mathbf{x}, \lambda) = \sum_{i=1}^{n} c_i(x_i) + \lambda \left(D - \sum_{i=1}^{n} x_i\right) \tag{5.15}$$

can be rewritten as

$$L(\mathbf{x}, \lambda) = -\sum_{i=1}^{n} \big(\lambda x_i - c_i(x_i)\big) + \lambda D \tag{5.16}$$

The ith term in the sum is the net profit of plant i if its output is valued at λ. Therefore, if the company undertakes to buy electricity from its plants at the price λ and instructs each plant manager to produce so as to maximize the plant's net profit, each manager will be induced to choose an output level which maximizes the profit of the company as a whole. This is the case whether the price λ is the market price at which the company can buy electricity from external suppliers or the shadow price determined by the need to satisfy the total demand D. In this way, the shadow price λ can be used to decentralize the production decision.

5.30 The Lagrangean for this problem is

$$L(x_1, x_2, \lambda_1, \lambda_2) = x_1 x_2 - \lambda_1(x_1^2 + 2x_2^2 - 3) - \lambda_2(2x_1^2 + x_2^2 - 3)$$

The first-order conditions for stationarity

$$\begin{aligned} D_{x_1} L &= x_2 - 2\lambda_1 x_1 - 4\lambda_2 x_1 = 0 \\ D_{x_2} L &= x_1 - 4\lambda_1 x_2 - 2\lambda_2 x_2 = 0 \end{aligned}$$

can be written as

$$x_2 = 2(\lambda_1 + 2\lambda_2)x_1 \tag{5.17}$$

$$x_1 = 2(2\lambda_1 + \lambda_2)x_2 \tag{5.18}$$

which must be satisfied along with the complementary slackness conditions

$$\begin{aligned} &x_1^2 + 2x_2^2 - 3 \leq 0 \qquad \lambda_1 \geq 0 \qquad \lambda_1(x_1^2 + 2x_2^2 - 3) = 0 \\ &2x_1^2 + x_2^2 - 3 \leq 0 \qquad \lambda_2 \geq 0 \qquad \lambda_2(2x_1^2 + x_2^2 - 3) = 0 \end{aligned}$$

First suppose that both constraints are slack so that $\lambda_1 = \lambda_2 = 0$. Then the first-order conditions (5.17) and (5.18) imply that $x_1 = x_2 = 0$. $(0,0)$ satisfies the Kuhn-Tucker conditions. Next suppose that the first constraint is binding while the second constraint is slack ($\lambda_2 = 0$). The first-order conditions (5.17) and (5.18) have two solutions, $x_1 = \sqrt{3/2}$, $x_2 = \sqrt{3}/2$, $\lambda = 1/(2\sqrt{2})$ and $x_1 = -\sqrt{3/2}$, $x_2 = -\sqrt{3}/2$, $\lambda = 1/(2\sqrt{2})$, but these violate the second constraint. Similarly, there is no solution in which the first constraint is slack and the second constraint binding. Finally, assume that the both constraints are binding. This implies that $x_1 = x_2 = 1$ or $x_1 = x_2 = -1$, which points satisfy the first-order conditions (5.17) and (5.18) with $\lambda_1 = \lambda_2 = 1/6$.

We conclude that three points satisfy the Kuhn-Tucker conditions, namely $(0,0)$, $(1,1)$ and $(-1,-1)$. Noting the objective function, we observe that $(0,0)$ in fact minimizes the objective. We conclude that there are two local maxima, $(1,1)$ and $(-1,-1)$, both of which achieve the same level of the objective function.

5.31 Dividing the first-order conditions, we obtain

$$\frac{D_k R(k,l)}{D_l R(k,l)} = \frac{r}{w} - \frac{\lambda(s-r)}{(1-\lambda)w}$$

Using the revenue function

$$R(k,l) = p(f(k,l))f(k,l)$$

the marginal revenue products of capital and labor are

$$\begin{aligned} D_k R(k,l) &= D_y p(y) D_k f(k,l) \\ D_l R(k,l) &= D_y p(y) D_l f(k,l) \end{aligned}$$

so that their ratio is equal to the ratio of the marginal products

$$\frac{D_k R(k,l)}{D_l R(k,l)} = \frac{D_k f(k,l)}{D_l f(k,l}$$

The necessary condition for optimality can be expressed as

$$\frac{D_k f(k,l)}{D_l f(k,l)} = \frac{r}{w} - \frac{\lambda}{1-\lambda}\frac{s-r}{w}$$

whereas the necessary condition for cost minimization is (Example 5.16)

$$\frac{D_k f(k,l)}{D_l f(k,l)} = \frac{r}{w}$$

The regulated firm does not use the cost-minimizing combination of inputs.

5.32 The general constrained optimization problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}$$

can be transformed into an equivalent equality constrained problem

$$\max_{\mathbf{x},\mathbf{s}} f(\mathbf{x})$$
$$\text{subject to } \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{0} \text{ and } \ \mathbf{s} \geq \mathbf{0}$$

through the addition of nonnegative slack variables $\mathbf{s}$. Letting $\hat{\mathbf{g}}(\mathbf{x},\mathbf{s}) = \mathbf{g}(\mathbf{x}) + \mathbf{s}$, the first-order conditions a local optimum are (Exercise 5.16)

$$D_{\mathbf{x}} f(\mathbf{x}^*) = \sum \lambda_j D_{\mathbf{x}} \hat{g}_j(\mathbf{x}^*, \mathbf{s}^*) = \sum \lambda_j D_{\mathbf{x}} g_j(\mathbf{x}^*)$$
$$\mathbf{0} = D_{\mathbf{s}} f(\mathbf{x}^*) \leq \sum \lambda_j D_{\mathbf{s}} \hat{g}_j(\mathbf{x}, \mathbf{s}) = \boldsymbol{\lambda} \tag{5.19}$$
$$\mathbf{s} \geq \mathbf{0} \qquad \boldsymbol{\lambda}^T \mathbf{s} = 0 \tag{5.20}$$

Condition (5.19) implies that $\lambda_j \geq 0$ for every j. Furthermore, rewriting the constraint as

$$\mathbf{s} = -\mathbf{g}(\mathbf{x})$$

the complementary slackness condition (5.20) becomes

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0} \qquad \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = 0$$

This establishes the necessary conditions of Theorem 5.3.

5.33 The equality constrained maximization problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}$$

is equivalent to the problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}$$
$$-\mathbf{g}(\mathbf{x}) \leq -\mathbf{0}$$

By the Kuhn-Tucker theorem (Theorem 5.3), there exists nonnegative multipliers $\lambda_1^+, \lambda_2^+, \dots, \lambda_m^+$ and $\lambda_1^-, \lambda_2^-, \dots, \lambda_m^-$ such that

$$Df(\mathbf{x}^*) = \sum \lambda_j^+ Dg_j[\mathbf{x}^*] - \sum \lambda_j^- Dg_j[\mathbf{x}^*] = \mathbf{0} \tag{5.21}$$

with

$$\lambda_j^+ g_j(\mathbf{x}) = \mathbf{0} \text{ and } \lambda_j^- g_j(\mathbf{x}) = \mathbf{0} \qquad j = 1, 2, \dots, m$$

Defining $\lambda_j = \lambda_j^+ - \lambda_j^-$, (5.21) can be written as

$$Df(\mathbf{x}^*) = \sum \lambda_j Dg_j[\mathbf{x}^*]$$

which is the first-order condition for an equality constrained problem. Furthermore, if $\mathbf{x}^*$ satisfies the inequality constraints

$$g(\mathbf{x}^*) \leq \mathbf{0} \text{ and } g(\mathbf{x}^*) \geq \mathbf{0}$$

it satisfies the equality

$$g(\mathbf{x}^*) = \mathbf{0}$$

5.34 Suppose that $\mathbf{x}^*$ solves the problem

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \text{ subject to } A\mathbf{x} \leq \mathbf{0}$$

with Lagrangean

$$L = \mathbf{c}^T\mathbf{x} - \boldsymbol{\lambda}^T A\mathbf{x}$$

Then there exists $\boldsymbol{\lambda} \geq 0$ such that

$$D_{\mathbf{x}}L = \mathbf{c}^T - \boldsymbol{\lambda}^T A = \mathbf{0}$$

that is, $A^T\boldsymbol{\lambda} = \mathbf{c}$. Conversely, if there is no solution, there exists $\mathbf{x}$ such that $A\mathbf{x} \leq \mathbf{0}$ and

$$\mathbf{c}^T\mathbf{x} > \mathbf{c}^T\mathbf{0} = 0$$

5.35 There are two binding constraints at $(4, 0)$, namely

$$\begin{aligned} g(x_1, x_2) &= x_1 + x_2 \leq 4 \\ h(x_1, x_2) &= -x_2 \leq 0 \end{aligned}$$

with gradients

$$\begin{aligned} \nabla g(4, 0) &= (1, 1) \\ \nabla h(4, 0) &= (0, 1) \end{aligned}$$

which are linearly independent. Therefore the binding constraints are regular at $(0, 4)$.

5.36 The Lagrangean for this problem is

$$L(\mathbf{x}, \lambda) = u(\mathbf{x}) - \lambda(\mathbf{p}^T\mathbf{x} - m)$$

and the first-order (Kuhn-Tucker) conditions are (Corollary 5.3.2)

$$\begin{array}{llll} D_{x_i}L[\mathbf{x}^*, \lambda] = D_{x_i}u[\mathbf{x}^*] - \lambda p_i \leq \mathbf{0} & \mathbf{x}_i^* \geq 0 & x_i^*(D_{x_i}u[\mathbf{x}^*] - \lambda p_i) = 0 & (5.22) \\ \mathbf{p}^T\mathbf{x}^* \leq m & \lambda \geq \mathbf{0} & \lambda(\mathbf{p}^T\mathbf{x}^* - m) = 0 & (5.23) \end{array}$$

for every good $i = 1, 2, \dots, m$. Two cases must be distinguished.

Case 1 $\boxed{\lambda > 0}$ This implies that $\mathbf{p}^T\mathbf{x} = m$, the consumer spends all her income. Condition (5.22) implies

$$D_{x_i}u[\mathbf{x}^*] \leq \lambda p_i \text{ for every } i \text{ with } D_{x_i}u[\mathbf{x}^*] = \lambda p_i \text{ for every } i \text{ for which } x_i > 0$$

This case was analyzed in Example 5.17.

Case 2 $\boxed{\lambda = 0}$ This allows the possibility that the consumer does not spend all her income. Substituting $\lambda = 0$ in (5.22) we have $D_{x_i}u[\mathbf{x}^*] = 0$ for every i. At the optimal consumption bundle $\mathbf{x}^*$, the marginal utility of every good is zero. The consumer is satiated, that is no additional consumption can increase satisfaction. This case was analyzed in Example 5.31.

In summary, at the optimal consumption bundle $\mathbf{x}^*$, either

- the consumer is satiated ($D_{x_i}u[\mathbf{x}^*] = 0$ for every i) or
- the consumer consumes only those goods whose marginal utility exceeds the threshold $D_{x_i}u[\mathbf{x}^*] \geq \lambda p_i$ and adjusts consumption so that the marginal utility is proportional to price for all consumed goods.

5.37 Assume $\mathbf{x} \in D(\mathbf{x}^*)$. Then there exists $\bar{\alpha} \in \Re$ such that $\mathbf{x}^* + \alpha\mathbf{x} \in S$ for every $0 \leq \alpha \leq \bar{\alpha}$. Define $g \in F([0, \bar{\alpha}])$ by $g(\alpha) = f(\mathbf{x}^* + \alpha\mathbf{x})$. If $\mathbf{x}^*$ is a local maximum, g has a local maximum at 0, and therefore $g'(0) \leq 0$ (Theorem 5.1). By the chain rule (Exercise 4.22), this implies

$$g'(0) = Df[\mathbf{x}^*](\mathbf{x}) \leq 0$$

and therefore $\mathbf{x} \notin H^+(\mathbf{x}^*)$.

5.38 If $\mathbf{x}$ is a tangent vector, so is $\beta\mathbf{x}$ for any nonnegative β (replace $1/\alpha^k$ by β/α^k in the preceding definition. Also, trivially, $\mathbf{x} = \mathbf{0}$ is a tangent vector (with $\mathbf{x}^k = \mathbf{x}^*$ and $\alpha_k = 1$ for all k). The set T of all vectors tangent to S at $\mathbf{x}^*$ is therefore a nonempty *cone*, which is called the *cone of tangents* to S at $\mathbf{x}^*$.

To show that T is closed, let $\mathbf{x}^n$ be a sequence in T converging to some $\mathbf{x} \in \Re^n$. We need to show that $\mathbf{x} \in T$. Since $\mathbf{x}^n \in T$, there exist feasible points $\mathbf{x}^{mn} \in S$ and α^{mn} such that

$$(\mathbf{x}^{mn} - \mathbf{x}^*)/\alpha^{mn} \to \mathbf{x}^n \text{ as } m \to \infty$$

For any N choose n such that

$$\|\mathbf{x}^n - \mathbf{x}\| \leq \frac{1}{2}N$$

and then choose m such that

$$\|\mathbf{x}^{mn} - \mathbf{x}^*\| \leq N \text{ and } \|(\mathbf{x}^{mn} - \mathbf{x}^*)/\alpha^{mn} - \mathbf{x}^n\| \leq \frac{1}{2}N$$

Relabeling $\mathbf{x}^{mn}$ as $\mathbf{x}^N$ and α^{mn} as α^N we have we have constructed a sequence $\mathbf{x}^N$ in S such that

$$\left\|\mathbf{x}^N - \mathbf{x}^*\right\| \leq N$$

and

$$\left\|(\mathbf{x}^N - \mathbf{x}^*)/\alpha^N - \mathbf{x}\right\| \leq \left\|(\mathbf{x}^N - \mathbf{x}^*)/\alpha^N - \mathbf{x}^n\right\| + \|\mathbf{x}^n - \mathbf{x}\| \leq \frac{1}{1}N$$

Letting $N \to \infty$, $\mathbf{x}^N$ converges to $\mathbf{x}^*$ and $(\mathbf{x}^N - \mathbf{x}^*)/\alpha^N$ converges to $\mathbf{x}$, which proves that $\mathbf{x} \in T$ as required.

5.39 Assume $\mathbf{x} \in D(\mathbf{x}^*)$. That is, there exists $\bar{\alpha}$ such that $\mathbf{x}^* + \alpha\mathbf{x} \in S$ for every $\alpha \in [0, \bar{\alpha}]$. For $k = 1, 2, \dots$, let $\alpha_k = \bar{\alpha}/k$. Then $\mathbf{x}^k = \mathbf{x}^* + \alpha_k\mathbf{x} \in S$, $\mathbf{x}^k \to \mathbf{x}^*$ and $(\mathbf{x}^k - \mathbf{x}^*)/\alpha_k = (\mathbf{x}^* + \alpha_k\mathbf{x} - \mathbf{x}^*)/\alpha_k = \mathbf{x}$. Therefore, $\mathbf{x} \in T(\mathbf{x}^*)$.

5.40 Let $\mathbf{dx} \in T(\mathbf{x}^*)$. Then there exists a feasible sequence $\{\mathbf{x}^k\}$ converging to $\mathbf{x}^*$ and a sequence $\{\alpha_k\}$ of nonnegative scalars such that the sequence $\{(\mathbf{x}^k - \mathbf{x}^*)/\alpha_k\}$ converges to $\mathbf{dx}$. For any $j \in B(\mathbf{x}^*)$, $g_j(\mathbf{x}^*) = 0$ and

$$g_j(\mathbf{x}^k) = Dg_j[\mathbf{x}^*](\mathbf{x}^k - \mathbf{x}^*) + \eta_j \left\|\mathbf{x}^k - \mathbf{x}^*\right\|$$

where $\eta_j \to 0$ as $\mathbf{k} \to \infty$. This implies

$$\frac{1}{\alpha^k} g_j(\mathbf{x}^k) = \frac{1}{\alpha^k} Dg_j[\mathbf{x}^*](\mathbf{x}^k - \mathbf{x}^*) + \eta_j \left\|(\mathbf{x}^k - \mathbf{x}^*)/\alpha^k\right\|$$

Since $\mathbf{x}^k$ is feasible

$$\frac{1}{\alpha^k} g_j(\mathbf{x}^k) \leq 0$$

and therefore

$$Dg_j[\mathbf{x}^*]((\mathbf{x}^k - \mathbf{x}^*)/\alpha^k) + \eta^i \left\|(\mathbf{x}^k - \mathbf{x}^*)/\alpha^k\right\| \leq \mathbf{0}$$

Letting $k \to \infty$ we conclude that

$$Dg_j[\mathbf{x}^*](\mathbf{dx}) \leq 0$$

That is, $\mathbf{dx} \in L$.

5.41 $L^0 \subseteq L^1$ by definition. Assume $\mathbf{dx} \in L^1$. That is

$$Dg_j[\mathbf{x}^*](\mathbf{dx}) < 0 \qquad \text{for every } \; j \in B^N(\mathbf{x}^*) \tag{5.24}$$

$$Dg_j[\mathbf{x}^*](\mathbf{dx}) \leq 0 \qquad \text{for every } \; j \in B^C(\mathbf{x}^*) \tag{5.25}$$

where $B^C(\mathbf{x}^*) = B(\mathbf{x}^*) - B^N(\mathbf{x}^*)$ is the set of *concave* binding constraints at $\mathbf{x}^*$. By concavity (Exercise 4.67), (5.25) implies that

$$g_j(\mathbf{x}^* + \alpha\mathbf{dx}) \leq g_j(\mathbf{x}^*) = 0 \text{ for every } \; \alpha \geq 0 \text{ and } \; j \in B^C(\mathbf{x}^*)$$

From (5.24) there exists some α_N such that

$$g_j(\mathbf{x}^* + \alpha\mathbf{dx}) < 0 \text{ for every } \; \alpha \in [0, \alpha_N] \text{ and } \; j \in B^N(\mathbf{x}^*)$$

Furthermore, since $g_j(\mathbf{x}^*) < 0$ for all $j \in S(\mathbf{x}^*)$, there exists some $\alpha_S > 0$ such that

$$g_j(\mathbf{x}^* + \alpha\mathbf{dx}) < 0 \text{ for every } \; \alpha \in [0, \alpha_S] \text{ and } \; j \in S(\mathbf{x}^*)$$

Setting $\bar{\alpha} = \min\{\alpha_N, \alpha_S\}$ we have

$$g_j(\mathbf{x}^* + \alpha\mathbf{dx}) \leq 0 \text{ for every } \; \alpha \in [0, \bar{\alpha}] \text{ and } \; j = 1, 2, \dots, m$$

or

$$\mathbf{x}^* + \alpha\mathbf{dx} \in G = \{\, \mathbf{x} : g_j(\mathbf{x}) \leq 0, \; j = 1, 2, \dots, m \,\} \text{ for every } \; \alpha \in [0, \bar{\alpha}]$$

Therefore $\mathbf{dx} \in D$. We have previously shown (Exercises 5.39 and 5.40) that $D \subset T \subset L$.

5.42 Assume that $\mathbf{g}$ satisfies the Quasiconvex CQ condition at $\mathbf{x}^*$. That is, for every $j \in B(\mathbf{x}^*)$, g_j is quasiconvex, $\nabla g_j(\mathbf{x}^*) \neq \mathbf{0}$ and there exists $\hat{\mathbf{x}}$ such that $g_j(\hat{\mathbf{x}}) < 0$. Consider the perturbation $\mathbf{dx} = \hat{\mathbf{x}} - \mathbf{x}^*$. Quasiconvexity and regularity implies that for every binding constraint $j \in B(\mathbf{x}^*)$ (Exercises 4.74 and 4.75)

$$g_j(\hat{\mathbf{x}}) < g_j(\mathbf{x}^*) \implies \nabla g_j(\mathbf{x}^*)^T(\hat{\mathbf{x}} - \mathbf{x}^*) = \nabla g_j(\mathbf{x}^*)^T \mathbf{dx} < 0$$

That is

$$Dg_j[\mathbf{x}^*](\mathbf{dx}) < \mathbf{0}$$

Therefore, $\mathbf{dx} \in L^0(\mathbf{x}^*) \neq \emptyset$ and $\mathbf{g}$ satisfies the Cottle constraint qualification condition.

5.43 If the binding constraints $B(\mathbf{x}^*)$ are regular at $\mathbf{x}^*$, their gradients are linearly independent. That is, there exists no $\lambda_j \neq 0, j \in B(\mathbf{x}^*)$ such that

$$\sum_{j \in B(\mathbf{x}^*)} \lambda_j \nabla g_j[\mathbf{x}^*] = \mathbf{0}$$

By Gordan's theorem (Exercise 3.239), there exists $\mathbf{dx} \in \Re^n$ such that

$$\nabla g_j[\mathbf{x}^*]^T \mathbf{dx} < 0 \text{ for every } \ j \in B(\mathbf{x}^*)$$

Therefore $\mathbf{dx} \in L^0(\mathbf{x}^*) \neq \emptyset$.

5.44 If g_j concave, $B^N(\mathbf{x}^*) = \emptyset$, and AHUCQ is trivially satisfied (with $\mathbf{dx} = \mathbf{0} \in L^1$).

For every j, let

$$S_j = \{\, \mathbf{dx} : Dg_j[\mathbf{x}^*](\mathbf{dx}) < 0 \,\}$$

Then

$$L^1(\mathbf{x}^*) = \left(\bigcap_{i \in B^N(\mathbf{x}^*)} S_i \right) \bigcap \left(\bigcap_{i \in B^C(\mathbf{x}^*)} \overline{S_i} \right)$$

where $B^C(\mathbf{x}^*)$ and $B^N(\mathbf{x}^*)$ are respectively the concave and nonconcave constraints binding at $\mathbf{x}^*$. If g_j satisfies the AHUCQ condition, $L^1(\mathbf{x}^*) \neq \emptyset$ and Exercise 1.219 implies that

$$\overline{L^1} = \left(\bigcap_{i \in B^N(\mathbf{x}^*)} \overline{S_i} \right) \bigcap \left(\bigcap_{i \in B^C(\mathbf{x}^*)} \overline{S_i} \right)$$

Now

$$\overline{S_i} = \{ \ \mathbf{dx} : Dg_j[\mathbf{x}^*](\mathbf{dx}) \leq 0 \ \}$$

and therefore

$$\overline{L^1} = \bigcap_{j \in B(\mathbf{x}^*)} \overline{S_i} = L$$

Since (Exercise 5.41)

$$L^1 \subseteq T \subseteq L$$

and T is closed (Exercise 5.38), we have

$$L = \overline{L^1} \subseteq T \subseteq L$$

which implies that $T = L$.

5.45 For each $j = 1, 2, \ldots, m$, either

$\boxed{g_j(\mathbf{x}^*) < 0}$ which implies that $\lambda_j = 0$ and therefore $\lambda_j Dg_j[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) = 0$

or

$\boxed{g_j(\mathbf{x}^*) = 0}$ Since g_j is quasiconvex and $g_j(\mathbf{x}) \leq 0 = g(\mathbf{x}^*)$, Exercise 4.73 implies that $Dg_j[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) \leq 0$. Since $\lambda_j \geq 0$, this implies that $\lambda_j Dg_j[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) \leq 0$.

We have shown that for every j, $\lambda_j Dg_j[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) \leq 0$. The first-order condition implies that

$$Df[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) = \sum_j \lambda_j Dg_j[\mathbf{x}^*](\mathbf{x} - \mathbf{x}^*) \leq 0$$

If

$$\nabla f(\mathbf{x}^*) \leq \sum \lambda_j \nabla g_j(\mathbf{x}^*) \qquad \mathbf{x}^* \geq 0 \qquad \Big(\nabla f(\mathbf{x}^*) - \lambda_j \nabla g_j(\mathbf{x}^*)\Big)^T \mathbf{x}^* = 0$$

The first-order conditions imply that for every $\mathbf{x} \in G$, $\mathbf{x} \geq 0$ and

$$\Big(\nabla f(\mathbf{x}^*) - \lambda_j \nabla g_j(\mathbf{x}^*)\Big)^T \mathbf{x} \leq 0$$

and therefore

$$\Big(\nabla f(\mathbf{x}^*) - \lambda_j \nabla g_j(\mathbf{x}^*)\Big)^T (\mathbf{x} - \mathbf{x}^*) \leq 0$$

or

$$\nabla f(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \leq \sum \lambda_j \nabla g_j(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \leq 0$$

5.46 Assuming $x_d = x_d = 0$, the constraints become

$$\begin{aligned} 2x_c &\leq 30 \\ 2x_c &\leq 25 \\ x_c &\leq 20 \end{aligned}$$

The first and third conditions are redundant, which implies that $\lambda_f = \lambda_m = 0$. Complementary slackness requires that, if $x_c > 0$,

$$D_{x_c} L = 1 - 2\lambda_f - 2\lambda_l - \lambda_m = 0$$

or $\lambda_l = \frac{1}{2}$. Evaluating the Lagrangean at $(0, 1/2, 0)$ yields

$$\begin{aligned} L\Big(\mathbf{x}, \Big(0, \frac{1}{2}, 0\Big)\Big) &= 3x_b + x_c + 3x_d \\ &\quad - \frac{1}{2}(x_b + 2x_c + 3x_d - 25) \\ &= \frac{25}{2} + \frac{5}{2}x_b + \frac{3}{2}x_d \end{aligned}$$

This basic feasible solution is clearly not optimal, since profit would be increased by increasing either x_b or x_d.

Following the hint, we allow $x_d > 0$, retaining the assumption that $x_b = 0$. We must be alert to the possibility that $x_c = 0$. With $x_b = 0$, the constraints become

$$\begin{aligned} 2x_c + x_d &\leq 30 \\ 2x_c + 3x_d &\leq 25 \\ x_c + x_d &\leq 20 \end{aligned}$$

The first constraint is redundant, which implies that $\lambda_f = 0$. If $x_d > 0$, complementary slackness requires that

$$D_{x_d} L = 3 - 3\lambda_l - \lambda_m = 0$$

or

$$\lambda_m = 3(1 - \lambda_l) \tag{5.26}$$

The requirement that $\lambda_m \geq 0$ implies that $\lambda_l \leq 1$. Substituting (5.26) in the second first-order condition

$$D_{x_c} L = 1 - 2\lambda_l - \lambda_m = 1 - 2\lambda_l - 3(1 - \lambda_l) = -2 + \lambda_l$$

implies that

$$D_{x_c} L = -2 + \lambda_l < 0 \qquad \text{for every} \quad \lambda_l \leq 1$$

Complementary slackness then requires implies that $x_c = 0$.

The constraints now become

$$\begin{aligned} x_d &\leq 30 \\ 3x_d &\leq 25 \\ x_d &\leq 20 \end{aligned}$$

The first and third are redundant, so that λ_f and $\lambda_m = 0$. Equation (5.26) implies that $\lambda_l = 1$.

Evaluating the Lagrangean at this point ($\lambda = 0, 1, 0$), we have

$$\begin{aligned} L(x, (0,1,0)) &= 3x_b + x_c + 3x_d \\ &\quad - (x_b + 2x_c + 3x_d - 25) \\ &= 25 + 2x_b - x_c \end{aligned}$$

Clearly this is not an optimal solution, An increase in x_b is indicated. This leads us to the hypothesis $x_b > 0$, $x_d > 0$, $x_c = 0$ which was evaluated in the text, and in fact lead to the optimal solution.

5.47 If we ignore the hint and consider solutions with $x_b > 0$, $x_c \geq 0$, $x_d = 0$, the constraints become

$$\begin{aligned} 2x_b + 2x_c &\leq 30 \\ x_b + 2x_c &\leq 25 \\ 2x_b + x_c &\leq 20 \end{aligned}$$

These three constraints are linearly dependent, so that any one of them is redundant and can be eliminated. For example, $3/2$ times the first constraint is equal to the sum of the second and third constraints. The feasible solution $x_b = 0$, $x_c = 5$, $x_d = 10$, where the constraints are linearly dependent, is known as a *degenerate* solution. Degeneracy is a significant feature of linear programming, allowing the theoretical possibility of a breakdown in the simplex algorithm. Fortunately, such breakdown seems very rare in practice. Degeneracy at the optimal solution indicates multiple optima.

One way to proceed in this example is to arbitrarily designate one constraint as redundant, assuming the corresponding multiplier is zero. Arbitrarily choosing $\lambda_m = 0$ and proceeding as before, complementary slackness ($x_d > 0$) requires that

$$D_{x_d} L = 3 - 2\lambda_f - \lambda_l = 0$$

or

$$\lambda_l = 3 - 2\lambda_f \tag{5.27}$$

Nonnegativity of λ_l implies that $\lambda_f \leq \frac{3}{2}$.

Substituting (5.27) in the second first-order condition yields

$$\begin{aligned} D_{x_c} L &= 1 - 2\lambda_f - 2\lambda_l \\ &= 1 - 2\lambda_f - 2(3 - 2\lambda_f) \\ &= -5 + 2\lambda_f < 0 \text{ for every } \ \lambda_f \leq \frac{3}{2} \end{aligned}$$

Complementary slackness therefore implies that $x_c = 0$, which takes us back to the starting point of the presentation in the text, where $x_b > 0$, $x_c = x_d = 0$.

5.48 Assume that $(\mathbf{c}_1, z_1)$ and $(\mathbf{c}_2, z_2)$ belong to B. That is

$$\begin{aligned} \mathbf{c}_1 \leq \mathbf{0} \qquad & z_1 \geq z^* \\ \mathbf{c}_2 \leq \mathbf{0} \qquad & z_2 \geq z^* \end{aligned}$$

For any $\alpha \in (0,1)$,

$$\bar{\mathbf{c}} = \alpha \mathbf{c}_1 + (1-\alpha)\mathbf{c}_2 \leq \mathbf{0} \qquad \bar{z} = \alpha z_1 + (1-\alpha) z_2 \leq z^*$$

and therefore $(\bar{\mathbf{c}}, \bar{z}) \in B$. This shows that B is convex. Let $\mathbf{1} = (1, 1, \dots, 1) \in \Re^m$. Then $(\mathbf{c} - \mathbf{1}, z + 1) \in \operatorname{int} B \neq \emptyset$. There B has a nonempty interior.

5.49 Let $(\mathbf{c}, z) \in \operatorname{int} B$. This implies that $\mathbf{c} < \mathbf{0}$ and $z > z^*$. Since v is monotone

$$v(\mathbf{c}) \leq v(\mathbf{0}) = \mathbf{z}^* < z$$

which implies that $(\mathbf{c}, z) \notin A$.

5.50 The linear functional L can be decomposed into separate components, so that there exists (Exercise 3.47) $\varphi \in Y^*$ and $\alpha \in \Re$ such that

$$L(\mathbf{c}, z) = \alpha z - \varphi(\mathbf{c})$$

Assuming $Y \subseteq \Re^m$, there exists (Proposition 3.4) $\boldsymbol{\lambda} \in \Re^m$ such that $\varphi(\mathbf{c}) = \boldsymbol{\lambda}^T \mathbf{c}$ and therefore

$$L(\mathbf{c}, z) = \alpha z - \boldsymbol{\lambda}^T \mathbf{c}$$

The point $(\mathbf{0}, z^* + 1)$ belongs to B. Therefore, by (5.75),

$$L(\mathbf{0}, z^*) \leq L(\mathbf{0}, z^* + 1)$$

which implies that

$$\alpha z^* - \boldsymbol{\lambda}^T \mathbf{0} \leq \alpha(z^* + 1) - \boldsymbol{\lambda}^T \mathbf{0}$$

or $\alpha \geq 0$. Similarly, let $\{ \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m \}$ denote the standard basis for $\Re^m$ (Example 1.79). For any $j = 1, 2, \ldots, m$, the point $(\mathbf{0} - \mathbf{e}_j, z^*)$ (which corresponds to decreasing resource j by one unit) belongs to B and therefore (from (5.75))

$$z^* - \boldsymbol{\lambda}^T(\mathbf{0} - \mathbf{e}_j) = z^* + \lambda_j \geq z^* - \boldsymbol{\lambda}^T \mathbf{0} = z^*$$

which implies that $\lambda_j \geq 0$.

5.51 Let

$$\hat{\mathbf{c}} = g(\hat{\mathbf{x}}) < \mathbf{0} \text{ and } \hat{z} = f(\hat{\mathbf{x}})$$

Suppose $\alpha = 0$. Then, since L is nonzero, at least one component of $\boldsymbol{\lambda}$ must be nonzero. That is, $\boldsymbol{\lambda} \gneq \mathbf{0}$ and therefore

$$\boldsymbol{\lambda}^T \hat{c} < 0 \tag{5.28}$$

But $(\hat{\mathbf{c}}, \hat{z}) \in A$ and (5.74) implies

$$\alpha \hat{z} - \boldsymbol{\lambda}^T \hat{\mathbf{c}} \leq \alpha z^* - \boldsymbol{\lambda}^T \mathbf{0}$$

and therefore $\alpha = 0$ implies

$$\boldsymbol{\lambda}^T \hat{c} \geq 0$$

contradicting (5.28). Therefore, we conclude that $\alpha > 0$.

5.52 The utility's optimization problem is

$$\max_{y, Y \geq 0} S(y, Y) = \sum_{i=1}^n \int_0^{y_i} (p_i(\tau) - c_i) d\tau - c_0 Y$$
$$\text{subject to } g_i(\mathbf{y}, Y) = y_i - Y \leq 0 \qquad i = 1, 2, \ldots, n$$

The demand independence assumption ensures that the objective function S is concave, since its Hessian

$$H_S = \begin{pmatrix} Dp_1 & 0 \ldots & 0 & 0 \\ 0 & Dp_2 & \ldots 0 & 0 \\ 0 & \ldots & Dp_n & 0 \\ 0 & \ldots & 0 & 0 \end{pmatrix}$$

is nonpositive definite (Exercise 3.96). The constraints are linear and hence convex. Moreover, there exists a point $(\mathbf{0}, 1)$ such that for every $i = 1, 2, \ldots, n$

$$g_i(\mathbf{0}, 1) = 0 - 1 < 0$$

Therefore the problem satisfies the conditions of Theorem 5.6. The optimal solution $(\mathbf{y}^*, Y^*)$ satisfies the Kuhn-Tucker conditions, that is there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that for every period $i = 1, 2, \ldots, n$

$$\begin{aligned} D_{y_i} L = p_i(y_i) - c_i - \lambda_i &\leq 0 & y_i &\geq 0 & y_i(p_i(y_i) - c_i - \lambda_i) &= 0 \\ y_i &\leq Y & \lambda_i &\geq 0 & \lambda(Y - y_i) &= 0 \end{aligned} \tag{5.29}$$

and that capacity be chosen such that

$$D_Y L = c_0 - \sum_{i=1}^n \lambda_i \le 0 \qquad Y \ge 0 \qquad Y\left(c_0 - \sum_{i=1}^n \lambda_i\right) = 0 \tag{5.30}$$

where L is the Lagrangean

$$L(y, Y, \lambda) = \sum_{i=1}^n \int_0^{y_i} (p_i(\tau) - c_i)d\tau - c_0 Y - \sum_{i=1}^n \lambda_i (y_i - Y)$$

In off-peak periods ($y_i < Y$), complementary slackness requires that $\lambda_i = 0$ and therefore from (5.29)

$$p_i(y_i) = c_i$$

assuming $y_i > 0$. In peak periods ($y_i = Y$)

$$p_i(y_i) = c_i + \lambda_i$$

We conclude that it is optimal to price at marginal cost in off-peak periods and charge a premium during peak periods. Furthermore, (5.30) implies that the total premium is equal to the marginal capacity cost

$$\sum_{i=1}^n \lambda_i = c_0$$

Furthermore, note that

$$\begin{aligned}
\sum_{i=1}^n \lambda_i y_i &= \sum_{\text{Peak}} \lambda_i y_i + \sum_{\text{Off-peak}} \lambda_i y_i \\
&= \sum_{y_i = Y} \lambda_i y_i + \sum_{\lambda_i = 0} \lambda_i y_i \\
&= \sum_{y_i = Y} \lambda_i Y \\
&= \sum_{i=1}^n \lambda_i Y = c_0 Y
\end{aligned}$$

Therefore, the utility's total revenue is

$$\begin{aligned}
R(y, Y) &= \sum_{i=1}^n p_i(y_i) y_i \\
&= \sum_{i=1}^n (c_i + \lambda_i) y_i \\
&= \sum_{i=1}^n c_i y_i + \sum_{i=1}^n \lambda_i y_i \\
&= \sum_{i=1}^n c_i y_i + c_0 Y = c(y, Y)
\end{aligned}$$

Under the optimal pricing policy, revenue equals cost and the utility breaks even.