Chapter 1

1.1. $\mathbf{A} = \begin{bmatrix} 11 & -1 \\ 0 & 0 \\ -3 & 5 \end{bmatrix}$.

1.2. (a) **B** must be a 3 x 2 matrix.

(b) $\mathbf{B} = \begin{bmatrix} 10 & -2 \\ 10 & -2 \\ 10 & -2 \end{bmatrix}$.

(c) $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 21 & -3 \\ 10 & -2 \\ 7 & 3 \end{bmatrix}$.

1.3. Since $\mathbf{C} = \begin{bmatrix} -3 & 2 \\ 3 & -6 \\ 5 & -4 \end{bmatrix}$, $(\mathbf{A} + \mathbf{B}) - \mathbf{C} = \begin{bmatrix} 24 & -5 \\ 7 & 4 \\ 2 & 7 \end{bmatrix}$.

1.4. (a) (**D** + **E**) = [5 -4 3 2]; (**D** + **E**) + **F** = [11 1 -1 3].

(b) (**E** + **F**) = [7 6 -4 2]; **D** + (**E** + **F**) = [11 1 -1 3].

1.5. Since **A** has dimensions 2 x 3, **B** must be 3 x n in order to postmultiply **A** and **B** must be m x 2 to premultiply **A**. (a) **BA** is possible, **AB** is not. (b) Both **AB** and **BA** are possible. (c) Neither is possible. (d) As in (b), both are possible. (e) Neither is possible.

1.6. $\mathbf{AB} = \begin{bmatrix} 17 & -16 & 31 \\ 45 & -32 & 83 \end{bmatrix}$, $\mathbf{AC} = \begin{bmatrix} 12 & 9 & 12 \\ 28 & 37 & 44 \end{bmatrix}$, $\mathbf{AB} + \mathbf{AC} = \begin{bmatrix} 29 & -7 & 43 \\ 73 & 5 & 127 \end{bmatrix}$; $\mathbf{B} + \mathbf{C} = \begin{bmatrix} 2 & 8 & 10 \\ 9 & -5 & 11 \end{bmatrix}$, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \begin{bmatrix} 29 & -7 & 43 \\ 73 & 5 & 127 \end{bmatrix}$.

1.7. (a) $\mathbf{BC} = \begin{bmatrix} 100 \\ 200 \end{bmatrix}$, **A**(**BC**) = -700. (b) **AB** = [-26 -22], (**AB**)**C** = -700.

1.8. $\mathbf{A}_{11} = \begin{bmatrix} 1 & 2 \\ 4 & 6 \end{bmatrix}$, $\mathbf{A}_{12} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$, $\mathbf{A}_{21} = [1\ 0]$, $\mathbf{A}_{22} = [0]$; $\mathbf{B}_{11} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{B}_{12} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\mathbf{B}_{21} = [3]$, $\mathbf{B}_{22} = [4]$. So $\mathbf{A}_{11}\mathbf{B}_{11} = \begin{bmatrix} 5 \\ 16 \end{bmatrix}$, $\mathbf{A}_{12}\mathbf{B}_{21} = \begin{bmatrix} 15 \\ 21 \end{bmatrix}$, $\mathbf{A}_{21}\mathbf{B}_{1} = [1]$, $\mathbf{A}_{22}\mathbf{B}_{21} = [0]$, $\mathbf{A}_{11}\mathbf{B}_{12} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}$, $\mathbf{A}_{12}\mathbf{B}_{22} = \begin{bmatrix} 20 \\ 28 \end{bmatrix}$, $\mathbf{A}_{21}\mathbf{B}_{12} = [0]$ and $\mathbf{A}_{22}\mathbf{B}_{22} = [0]$. This means that

$$\mathbf{AB} = \left[\begin{array}{c|c} \begin{bmatrix} 5 \\ 16 \end{bmatrix} + \begin{bmatrix} 15 \\ 21 \end{bmatrix} & \begin{bmatrix} 4 \\ 12 \end{bmatrix} + \begin{bmatrix} 20 \\ 28 \end{bmatrix} \\ \hline [1] + [0] & [0] + [0] \end{array}\right] = \left[\begin{array}{c|c} 20 & 24 \\ 37 & 40 \\ \hline 1 & 0 \end{array}\right]$$

which is the same as the product **AB** obtained directly.

1.9. (a) -2. (b) 1. (c) -4.

1.10. (a) Columns (and rows) equal. (b) Column 3 = 3 x column 1. (c) Row of zeros. (d) Column 2 = 5 x column 3. (e) Row 1 = row 3.

1.11. (a) $|\mathbf{A}| = -1 \neq 0$, so $\mathbf{A}^{-1}$ exists and there is a unique solution. (b) Same as (a). (c) $|\mathbf{A}| = 0$; equation 2 is just four times equation 1, so this is one equation in two unknowns. (d) $|\mathbf{A}| = 0$; the left-hand side is the same as in (c), but now $b_2 = 25$, which is not four times b_1. These are two parallel lines in solution space.

1.12. $\mathbf{X} = \begin{bmatrix} 4 \\ 14 \end{bmatrix}$.

1.13. (a) $\mathbf{A}^{-1} = \begin{bmatrix} 4/11 & 1/11 \\ -3/11 & 1/22 \end{bmatrix}$. (b) $|\mathbf{B}| = 0$ so there is no inverse. (c) $\mathbf{C}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$. This illustrates the general rule for inverses of diagonal matrices that the inverse contains reciprocals of the original elements. (d) $\mathbf{D}^{-1} = \begin{bmatrix} 0.5 & 0.25 & -1 \\ 0 & -0.5 & 1 \\ -0.5 & -0.25 & 2 \end{bmatrix}$.

1.14. (a) $\rho(\mathbf{A}) = 1$. (b) $\rho(\mathbf{B}) = 2$. (c) $\rho(\mathbf{C}) = 1$. (d) $\rho(\mathbf{D}) = 1$. (e) $\rho(\mathbf{E}) = 2$. (f) $\rho(\mathbf{F}) = 3$. (g) $\rho(\mathbf{G}) = 2$.

1.15. Evaluate $|\mathbf{A}|$ down column j; evaluate $|\mathbf{A}'|$ across row j. The elements in column j of $\mathbf{A}$ are the same as the elements in row j of $\mathbf{A}'$. It is also easy to show that the cofactors of the elements in column j of $\mathbf{A}$ are the same as the cofactors of the elements in row j of $\mathbf{A}'$.

1.16. From (1-17), adj $\mathbf{A} = [\mathbf{A}_{ji}]$. Then (adj $\mathbf{A})' = [\mathbf{A}_{ij}]$; $\mathbf{A}' = [a_{ji}]$, so adj $\mathbf{A}' = [\mathbf{A}_{ij}]$ and (adj $\mathbf{A})' =$ adj $\mathbf{A}' = [\mathbf{A}_{ij}]$.

1.17. From (1-18), $\mathbf{A}^{-1} = (1/|\mathbf{A}|)(\text{adj } \mathbf{A})$. So $(\mathbf{A}^{-1})' = (1/|\mathbf{A}|)(\text{adj } \mathbf{A})'$. Also from (1-18), $(\mathbf{A}')^{-1} = (1/|\mathbf{A}'|)(\text{adj } \mathbf{A}')$. Since (adj $\mathbf{A}$) = (adj $\mathbf{A}'$) (Problem 1.16) and $|\mathbf{A}| = |\mathbf{A}'|$ (Problem 1.15), $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$.

1.18. For an n x n matrix $\mathbf{A}$, use elements from row i, $a_{i1}, \ldots, a_{in}$ and cofactors from row k ($\neq$ i), $\mathbf{A}_{k1}, \ldots, \mathbf{A}_{kn}$, giving

$$a_{i1}\mathbf{A}_{k1} + \ldots + a_{in}\mathbf{A}_{kn} \qquad (*)$$

Now imagine that row k in **A** is replaced by row i; call this matrix $\tilde{\mathbf{A}}$, so rows i and k in $\tilde{\mathbf{A}}$ are the same. From Property 4(a) of determinants, we know that $|\tilde{\mathbf{A}}| = 0$, because $\tilde{\mathbf{A}}$ has two equal rows. Find $|\tilde{\mathbf{A}}|$ by ordinary expansion across row k; this will be

$$|\tilde{\mathbf{A}}| = \alpha_{i1}A_{k1} + \dots + \alpha_{in}A_{kn} = 0$$

This is the same as (*), which must therefore also be zero.

1.19. Let **A** be an m x n matrix $\mathbf{A} = [\mathbf{A}_1 \vdots \dots \vdots \mathbf{A}_n]$. Then $\mathbf{A}' = \begin{bmatrix} (\mathbf{A}_1)' \\ \vdots \\ (\mathbf{A}_n)' \end{bmatrix}$ and $\mathbf{AA}' = [\mathbf{A}_1(\mathbf{A}_1)' + \dots + \mathbf{A}_n(\mathbf{A}_n)']$. Consider a typical element in this summation,

$$\mathbf{A}_j(\mathbf{A}_j)' = \begin{bmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{bmatrix} [\alpha_{1j} \dots \alpha_{nj}] = \begin{bmatrix} \alpha_{1j}\alpha_{1j} & \alpha_{1j}\alpha_{2j} & \dots & \alpha_{1j}\alpha_{nj} \\ \alpha_{2j}\alpha_{1j} & \alpha_{2j}\alpha_{2j} & \dots & \alpha_{2j}\alpha_{nj} \\ \vdots & \vdots & & \vdots \\ \alpha_{nj}\alpha_{1j} & \alpha_{nj}\alpha_{2j} & \dots & \alpha_{nj}\alpha_{nj} \end{bmatrix}.$$

This is symmetric, for any j = 1, ..., n, and so each of the products in $[\mathbf{A}_1(\mathbf{A}_1)' + \dots + \mathbf{A}_n(\mathbf{A}_n)']$ will be symmetric.

1.20. The element in row i and column j of **AB** is $({}_i\mathbf{A})(\mathbf{B}_j)$, where ${}_i\mathbf{A}$ = *row* i of **A**. Then the element in row j and column i of $(\mathbf{AB})'$ is $(\mathbf{B}_j)'({}_i\mathbf{A})'$. Row j of $\mathbf{B}'$ is $(\mathbf{B}_j)'$ and column i of $\mathbf{A}'$ is $({}_i\mathbf{A})'$, so the element in row j, column i of $\mathbf{B}'\mathbf{A}' = (\mathbf{B}_j)'({}_i\mathbf{A})'$. Since the element in row j, column i of $(\mathbf{AB})'$ is the same as the element in row j, column i of $\mathbf{B}'\mathbf{A}'$, the two matrices are equal.

1.21. Given (**AB**) and from the fundamental definition of an inverse, $(\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{I}$. Since **A** is nonsingular, $\mathbf{B}(\mathbf{AB})^{-1} = \mathbf{A}^{-1}$. Since **B** is also nonsingular, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

1.22. Suppose $\mathbf{A} = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$; then $\mathbf{A}^{-1} = \begin{bmatrix} -0.25 & -2 & 1.75 \\ 0.5 & -1 & 0.5 \\ -0.25 & 2 & -1.25 \end{bmatrix}$.

Let $\alpha = 2$; $(2\mathbf{A}) = \begin{bmatrix} 2 & 8 & 6 \\ 4 & 6 & 8 \\ 6 & 8 & 10 \end{bmatrix}$ and $(2\mathbf{A})^{-1} = \begin{bmatrix} -0.125 & -1 & 0.875 \\ 0.25 & -0.5 & 0.25 \\ -0.125 & 1 & -0.625 \end{bmatrix}$, which is clearly $(1/2)\mathbf{A}^{-1}$.

For the general case, $(\alpha\mathbf{A})^{-1} = \mathbf{A}^{-1}\alpha^{-1} = \alpha^{-1}\mathbf{A}^{-1}$ (since α is a scalar) $= (1/\alpha)\mathbf{A}^{-1}$ which is what was to be shown.

1.23. (a) Yes, since $|\mathbf{A}| \neq 0$. (b) Yes, for the same reason as in (a). (c) No, since the 4 x 4 matrix made up of these four columns is singular. (d) No, because there are too few vectors; three linearly independent vectors are needed for three dimensional vector space.

1.24. Using the **X′AX** form for these quadratic forms

(a) $\mathbf{A} = \begin{bmatrix} -1 & 0.5 & 0 \\ 0.5 & -1 & 1 \\ 0 & 1 & -4 \end{bmatrix}$. $|\mathbf{A}_1| = -1$, $|\mathbf{A}_2| = 0.75$, $|\mathbf{A}_3| = -2$, so the quadratic form is negative definite [from (1-40)].

(b) $\mathbf{A} = \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$. $|\mathbf{A}_1| = 2$, $|\mathbf{A}_2| = 3$, so the quadratic form is positive definite [from (1-40)].

(c) $\mathbf{A} = \begin{bmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{bmatrix}$. $|\mathbf{A}_1| = 1$, $|\mathbf{A}_2| = 0.75$, $|\mathbf{A}_3| = 0$, so this quadratic form is neither positive definite nor negative definite. However, $|\mathcal{A}_1(1)| = |\mathcal{A}_1(2)| = |\mathcal{A}_1(3)| = 1$; $|\mathcal{A}_2(1)| = |\mathcal{A}_2(2)| = |\mathcal{A}_2(3)| = 0.75$; $|\mathcal{A}_3(1)| = |\mathbf{A}_3| = 0$, so the quadratic form is positive semi-definite [from (1-41)].

Chapter 2

2.1. (a) $\rho(\mathbf{A}) = 1$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. Inconsistent system with no solution; the equations are parallel lines in solution space.

(b) $\rho(\mathbf{A}) = 2$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. Homogeneous equations, so the system is consistent. Since $\rho(\mathbf{A}) < 3$, there are multiple solutions. The two planes intersect in a line in solution space.

(c) $\rho(\mathbf{A}) = 2$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. Again, a consistent system. Any one equation is redundant; the remaining set of two equations in three unknowns has multiple solutions.

(d) $\rho(\mathbf{A}) = 2$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$. Inconsistent system; no solutions.

2.2. (a) There are two basic solutions, $\mathbf{X} = \begin{bmatrix} 0 \\ -10 \\ 5 \end{bmatrix}$ and $\mathbf{X} = \begin{bmatrix} 20 \\ -10 \\ 0 \end{bmatrix}$.

(b) Since $\rho(\mathbf{A}) = 2$ and $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$, any equation can be dropped, leaving a consistent system of two equations and three unknowns with at most C_2^3 basic solutions. They are (i) with $x_1 = 0$, $x_2 = 0.5$ and $x_3 = 1$. (ii) With $x_2 = 0$, $x_1 = -1$ and $x_3 = 2$. (iii) With $x_3 = 0$, $x_1 = x_2 = 1$.

(c) Here there are possibly as many as $C_2^4 = 6$ basic solutions. (i) Set $x_1 = x_2 = 0$. Then $x_3 = -10$, $x_4 = 10$. (ii) Set $x_1 = x_3 = 0$. Then $x_2 = 5$, $x_4 = 0$. (iii) Set $x_1 = x_4 = 0$. Then $x_2 = 5$, $x_3 = 0$. [Note that this is not distinguishably different from the solution in (ii).] (iv) Set $x_2 = x_3 = 0$. Then $x_1 = 10$, $x_4 = 0$. (v) Set $x_2 = x_4 = 0$. Then $x_1 = 10$, $x_3 = 0$. [This is the same as the solution in (iv).] (vi) Set $x_3 = x_4 = 0$. The coefficient matrix of the remaining 2 x 2 system is singular, so this possible basic solution does not exist.

2.3. (a) $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. This is a consistent system with a unique solution; but equation 3 = 3 x equation 1, so either can be dropped. Dropping equation 3, the solution is $x_1 = -19$, $x_2 = 11$. You will find the same solution if equation 1 is dropped.

(b) $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$. This is an inconsistent system. If any one equation is ignored, the remaining 2 x 2 system has a unique solution.

(c) Again, $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$. Inconsistent system; equations 1 and 3 are parallel, so solutions to a reduced system with equations 1 and 2 or with equations 2 and 3 could be found.

(d) $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. This is a consistent system; if *any* equation is dropped, the remaining system can be solved for a unique solution. [If equation 3 is dropped, this is just the system in part (a), above.]

2.4. (a) Since $|\mathbf{A}| = 0$, the answer is no.

(b) Here $|\mathbf{A}| \neq 0$, so there is a unique solution.

(c) Same as (b); there is a unique solution.

2.5. (a) $\rho(\mathbf{A}) = 2$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$ (= m and n). Inconsistent system. Equations 1 and 3 are parallel planes in three dimensional solution space; equation 2 cuts them.

(b) $\rho(\mathbf{A}) = 1$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 2 < m$. Inconsistent system with one redundant equation. Equations 1 and 2 describe the same plane; equation 3 is parallel to it.

(c) $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. Consistent system with a unique solution. But since $\rho(\mathbf{A}) < m$, there is one redundant equation.

(d) $\rho(\mathbf{A}) = 2$ (= n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$. Inconsistent system. The three equations intersect in three different points in two dimensional solution space.

(e) $\rho(\mathbf{A}) = 2$ (< n), $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$ (< m). This set of homogeneous equations has nontrivial, nonunique solutions. Since $\rho(\mathbf{A} \vdots \mathbf{B}) \neq m$, one [$m - \rho(\mathbf{A} \vdots \mathbf{B}) = 3 - 2$] equation is redundant. These are two intersecting planes in three dimensional solution space.

(f) $\rho(\mathbf{A}) = 1$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 1$. Consistent system with one redundant equation (equation 2 = 2 x equation 1). This is one plane in three dimensional solution space.

(g) $\rho(\mathbf{A}) = 1$, $\rho(\mathbf{A} \vdots \mathbf{B}) = 2$. Inconsistent system; two parallel planes.

2.6. (a) For a homogeneous system, c = 0. Since $|\mathbf{A}| = ab - 10$, the system will have multiple solutions when ab = 10, so that $|\mathbf{A}| = 0$.

(b) For a nonhomogeneous system, $c \neq 0$. Using Cramer's rule, $x_1 = -c/(ab - 10)$ and $x_2 = ac/(ab - 10)$, so $x_2 = 0$ if and only if $a = 0$.

2.7. (a) Here $|\mathbf{A}| = - a - b + 2$. So, if $a + b \neq 2$, $|\mathbf{A}| \neq 0$ and there will be a unique solution.

(b) The system will be inconsistent if $\rho(\mathbf{A}) < \rho(\mathbf{A}\,|\,\mathbf{B})$; that is, when $\rho(\mathbf{A}\,|\,\mathbf{B}) = 3$ and $\rho(\mathbf{A}) = 2$, as when $a + b = 2$. [This is the only possibility; $\rho(\mathbf{A}\,|\,\mathbf{B}) = 2$ and $\rho(\mathbf{A}) = 1$ is impossible since there is no choice of a and b that makes $\rho(\mathbf{A}) = 1$.] We will have $\rho(\mathbf{A}\,|\,\mathbf{B}) = 3$ if any of the following hold:

$\begin{vmatrix} 2 & a & 10 \\ b & -4 & -30 \\ -1 & 1 & -20 \end{vmatrix} = 3a + 2ab + b + 18 \neq 0$, or $\begin{vmatrix} 2 & -1 & 10 \\ b & 1 & -30 \\ -1 & 0 & -20 \end{vmatrix} = 2b + 6$ $\neq 0$ or $\begin{vmatrix} a & -1 & 10 \\ -4 & 1 & -30 \\ 1 & 0 & -20 \end{vmatrix} = 2a - 10 \neq 0$. Therefore, if $a \neq 5$ but $a + b = 2$, or if $b \neq -3$ but $a + b = 2$, the system will be inconsistent.

(c) There will be infinitely many solutions if the system is consistent and if $\rho(\mathbf{A})$ $[= \rho(\mathbf{A}\,|\,\mathbf{B})] < 3$; that is, if $a = 5$ and $b = -3$.

2.8. (a) There may be as many as $C_2^3 = 3$ basic solutions. However, letting $x_1 = 0$ produces a 2 x 2 system with a singular coefficient matrix. If $x_2 = 0$, then $x_1 = 2$ and $x_3 = 1$. If $x_3 = 0$, then $x_1 = 2$ and $x_2 = 2$.

(b) Any nonbasic solution is found by setting either x_2 or x_3 equal to some nonzero number and solving the remaining system. For example, if $x_2 = 5$, the equation system is

$$4x_1 + 6x_3 = -1$$
$$10x_1 + 4x_3 = 14$$

with solution $x_1 = 2$ and $x_3 = -1.5$; so the complete nonbasic solution is $\mathbf{X} = [2 \quad 5 \quad -1.5]'$

2.9. (a) Here $\rho(\mathbf{A}) = 2$ $(< n = 3)$, so there is no way that the system can have a unique solution.

(b) For there to be partial solutions, the system must be inconsistent. This means that $\rho(\mathbf{A}\,|\,\mathbf{B})$ must be 3; this will be true if, for example, $\mathbf{B} = [100 \quad 5 \quad 208]'$. Then partial solutions will be those that satisfy two of the three equations. Using just equations 1 and 2, a basic solution can be found by letting, say,

$x_3 = 0$. Then $x_1 = 0.8537$ and $x_2 = 0.4848$, so $\mathbf{X} = [0.8537 \;\; 0.4848 \;\; 0]'$ is a partial (and basic) solution to this system. Letting $x_3 = c$, for $c \neq 0$, would generate a nonbasic partial solution. Also, the right inverse for the 2 x 3 system of equations could be used to generate a partial solution.

2.10. (a) Since $\rho(\mathbf{A}) = 3$ ($< n = 4$), the kinds of solutions possible depend on **B** and in particular on $\rho(\mathbf{A}|\mathbf{B})$. If $\rho(\mathbf{A}|\mathbf{B}) = 4$, the system is inconsistent. The kinds of possible partial solutions are illustrated in Figures 2.5(f) and (g) for the $n = 3$ case. If $\rho(\mathbf{A}|\mathbf{B}) = 3$, there are multiple solutions. The analog in three dimensional solution space is illustrated by Figure 2.5(d).

(b) Here $\rho(\mathbf{A}) = 4$, so there will be a unique solution.

(c) As in (a), $\rho(\mathbf{A}) = 3$; now $\rho(\mathbf{A}|\mathbf{B}) = 3$, so there are multiple solutions.

(d) Again, $\rho(\mathbf{A}) = 3$ but now $\rho(\mathbf{A}|\mathbf{B}) = 4$, so the system is inconsistent.

2.11. (a) Two linear equations, one unknown; consistent system. Therefore either equation is redundant and there is a unique solution. Example: $2x_1 = 4$, $x_1 = 2$.

(b) Three linear equations, three unknowns; inconsistent system. Example: two parallel planes in three dimensional solution space, and a third plane that cuts through them.

(c) Four linear equations, four unknowns. The rank results are impossible, because either $\rho(\mathbf{A})$ and $\rho(\mathbf{A}|\mathbf{B})$ are equal or else $\rho(\mathbf{A}|\mathbf{B})$ is larger by 1. Conclusion: Computer or human error!

(d) Three equations, two unknowns; consistent system. Therefore there is one redundant equation; the unique solution can be found using any two of the three equations. Example: $x_1 + 0x_2 = 10$, $0x_1 + x_2 = 20$, $x_1 + x_2 = 30$.

(e) Two equations, two unknowns. Same conclusion as (c), since the smallest rank that a matrix can have is 1 (except for the trivial case of the null matrix, but if that were the coefficient matrix for an equation system it would mean that you had no equation system at all; try it).

2.12. Here $\rho(\mathbf{A}) = 2$ and $\rho(\mathbf{A} \vdots \mathbf{B}) = 3$. (You might notice that the first two equations define the same line; the second is just -7 times the first.) This is an inconsistent system with partial solutions. For equations 1 and 3, $\mathbf{X} = [0.4 \quad 0.4]'$; for equations 1 and 4, $\mathbf{X} = [0.5 \quad 1]'$; for equations 3 and 4, $\mathbf{X} = [0.5 \quad 0.5]'$.

2.13. (a) This results from the fact that the determinant of a triangular matrix is the product of the elements on the main diagonal, which is obvious in the 2 x 2 case and not difficult to show (from the general definition of a determinant) for the n x n case. For example, for an upper triangular matrix **A**, evaluate the determinant by expansion along the bottom row which contains all zeros except for a_{nn}. So $|\mathbf{A}| = a_{nn}\mathbf{A}_{nn}$. But by similar reasoning, $\mathbf{A}_{nn} = a_{n-1,n-1}\mathbf{A}_{n-1,n-1}$, etc. Eventually it will turn out that $|\mathbf{A}| = a_{nn}a_{n-1,n-1}\cdots a_{22}a_{11}$, which is what was to be shown.

(b) If you examine the general case for a 3 x 3 upper triangular matrix **A**, you will see that (adj **A**) must be upper triangular and hence $\mathbf{A}^{-1}$ will also be upper triangular. Then the same result follows for the 4 x 4, ..., n x n case, by examination of the structure of the cofactors in increasingly large cases.

(c) Let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 4 & 8 \end{bmatrix}$; then $\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ and $\mathbf{U} = \begin{bmatrix} 2 & 1 \\ 0 & 6 \end{bmatrix}$. Straightforward matrix multiplication confirms that $\mathbf{A} = \mathbf{LU}$ and that $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$.

For the general case, let $\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Starting with column 1, subtraction of (d/a) x row 1 from row 2 and subtraction of (g/a) x row 1 from row 3 produces

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ 0 & m_{33}/a & m_{32}/a \\ 0 & m_{23}/a & m_{22}/a \end{bmatrix}$$

where m_{ij} is the minor of a_{ij}. Next, subtraction of (m_{23}/m_{33}) x row 2 from row 3 produces

$$\mathbf{U} = \begin{bmatrix} a & b & c \\ 0 & m_{33}/a & m_{32}/a \\ 0 & 0 & (m_{22}/a) - (m_{23}m_{32}/m_{33}a) \end{bmatrix}$$

Recording these row multipliers in the lower part of a 3 x 3 identity matrix generates

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ d/a & 1 & 0 \\ g/a & m_{23}/m_{33} & 1 \end{bmatrix}$$

You can easily establish that $\mathbf{A} = \mathbf{LU}$. You can also find both $\mathbf{U}^{-1}$ and $\mathbf{L}^{-1}$ and then show that (i) $\mathbf{LL}^{-1} = \mathbf{I}$ and $\mathbf{UU}^{-1} = \mathbf{I}$ and (ii) $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$.

2.14. (a) $\mathbf{X} = \begin{bmatrix} 24 \\ 48 \\ 48 \end{bmatrix}$ (in millions of dollars). (b) $\mathbf{X} = \begin{bmatrix} 0.4\beta \\ 0.3\beta \\ 0.3\beta \end{bmatrix}$ (in millions of dollars).

2.15. (a) The equations are

$$\begin{aligned} x_{11} + x_{12} + x_{13} \qquad\qquad\qquad\qquad &= 10 \\ x_{21} + x_{22} + x_{23} &= 25 \\ x_{11} \qquad\quad + x_{21} \qquad\qquad &= 15 \\ x_{12} \qquad\quad + x_{22} \qquad &= 15 \\ x_{13} \qquad\quad + x_{23} &= 5 \end{aligned}$$

(b) In $\mathbf{AX} = \mathbf{B}$ form, this is

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} 10 \\ 25 \\ 15 \\ 15 \\ 5 \end{bmatrix}$$

(c) This is a set of 5 linear equations in 6 unknowns. The maximum number of basic solutions is $C_5^6 = 6$. However, since $\rho(\mathbf{A}\,|\,\mathbf{B}) = 4$ = the number of distinguishably different equations, this is really a system of 4 different linear equations in 6 unknowns. (You can see, for example, that equation 6 is a linear combination of equations 1 through 5; namely equation 6 = equation 1 + equation 2 + equation 3 - equation 4 - equation 5.) This means that there are a maximum of $C_4^6 = 15$ basic solutions. Here is the coefficient matrix, $\tilde{\mathbf{A}}$, that remains after equation 5 is deleted:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

(d) Let $\tilde{\mathbf{B}} = [10 \quad 25 \quad 15 \quad 15]'$, the original **B** with b_5 removed. You can easily verify that there are at least two 4 x 4 submatrices in $\tilde{\mathbf{A}}$ for which $\rho(\tilde{\mathbf{A}}) = \rho(\tilde{\mathbf{A}} \mid \tilde{\mathbf{B}})$ and so for which basic solutions exist.

2.16. Here is an example

$$\begin{aligned} x_1 + x_2 &= 10 \\ 2x_1 + 3x_2 &= 40 \\ 3x_1 + 3x_2 &= 30 \\ 4x_1 + 6x_2 &= 80 \\ 7x_1 + 7x_2 &= 70 \end{aligned}$$

In this case, $\rho(\mathbf{A}) = \rho(\mathbf{A} \mid \mathbf{B}) = 2$; equations 1, 3 and 5 define the same line; equations 2 and 4 define a second line in solution space. The unique solution is $\mathbf{X} = [-10 \quad 20]'$.

2.17. (a) $\mathbf{X} = \begin{bmatrix} 102 \\ -23 \end{bmatrix}$.

(b) Now the system is inconsistent.

(c) As in part (b), $\rho(\mathbf{A}) = 1$, since **A** is the same. The first element in **B** is now 200, so the question is, what value must b_2 have so that $\rho(\mathbf{A} \mid \mathbf{B}) = 1$ also, making the system consistent. The factor of proportionality between rows 1 and 2 of **A** is 200, so b_2 must be 40,000.

2.18. Sell 260 parcels of poorly developed land and buy 150 parcels of undeveloped land.

2.19. Increase type 1 projects by 54; decrease types 2 and 3 by 19 and 17, respectively.

2.20. Here $(\mathbf{A} \mid \mathbf{B}) = \begin{bmatrix} 1 & 1 & 1 & 10 \\ 2 & 2 & 2 & 20 \\ 7 & 3 & 3 & 10 \end{bmatrix}$. Subtracting twice row 1 from row 2 and seven times row 1 from row 3 (in order to produce a 0 in row 2, column 1 and in row 3, column 1), generates $\begin{bmatrix} 1 & 1 & 1 & 10 \\ 0 & 0 & 0 & 0 \\ 0 & -4 & -4 & -60 \end{bmatrix}$. Interchanging rows 2 and 3 and then dividing row 2 by -4 produces $\begin{bmatrix} 1 & 1 & 1 & 10 \\ 0 & 1 & 1 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. It is clear that equation 2 in the original set (now row 3) can be ignored; if you look at the original (**A**|**B**) matrix, it is clear that equations 1 and 2 are equal, so either can be ignored. We are left with two equations and three unknowns, for which basic solutions can be found. For example, setting $x_3 = 0$, you would have $x_2 = 15$ (from row 2 of the final matrix) and so $x_1 = -5$ (from row 1 of that matrix).

2.21. Here $(\mathbf{A}|\mathbf{B}) = \begin{bmatrix} 1 & 2 & 3 & 3 \\ 2 & 4 & 6 & 3 \\ 1 & 1 & 1 & 3 \end{bmatrix}$ and after two elementary row operations to generate zeros in positions a_{21} and a_{31}, we have $\begin{bmatrix} 1 & 2 & 3 & 3 \\ 0 & 0 & 0 & -3 \\ 0 & -1 & -2 & 0 \end{bmatrix}$. Dividing row 3 by -1 and interchanging rows 2 and 3 leads to $\begin{bmatrix} 1 & 2 & 3 & 3 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$. The inconsistency in this equation system is clear from row 3, where it is required that $0x_1 + 0x_2 + 0x_3 = -3$. (Compare with the final reduced matrix in Problem 2.20, where there was no inconsistency.)

2.22. Subtracting twice row 1 from row 2, to force a_{21} to zero, puts zeros in all of row 2. Interchanging rows 2 and 3, subtracting row 1 from (the new) row 2, and then dividing row 2 by -1 leads to $\begin{bmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & -2 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. As in Problem 2.20, the second equation (here now in row 3) can be ignored; in the original system, equations 1 and 2 are the same. So basic solutions can be found to two equations in four unknowns. For example, setting $x_3 = x_4 = 0$, $x_2 = 2$ (from row 2 of the final matrix) and so $x_1 = 0$ (from row 1 of that matrix). So this happens to be a degenerate basic solution.

2.23. Forcing a_{31} to zero (subtracting row 1 from row 3) generates $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$. Again, this shows that an equation, this time either the second or the third, is unnecessary (redundant), since these rows now contain exactly the same information. A solution is immediately seen to be $x_1 = x_2 = 2$. (The 2 x 3 submatrix containing the first two rows of the final matrix is in fact in reduced echelon form.)

2.24. After forcing α_{31} to zero (again, by subtracting row 1 from row 3), we have $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$, and it is clear that the requirements in rows 2 and 3 are inconsistent.

2.25. (a) $(\mathbf{A}|\mathbf{B}) = \begin{bmatrix} 3 & 2 & 7 \\ 2 & 5 & 12 \end{bmatrix}$ which can be reduced to $\begin{bmatrix} 1 & 0.667 & 2.333 \\ 0 & 1 & 2 \end{bmatrix}$ (echelon form). (b) The reduced matrix in (a) further reduces to $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$ (reduced echelon form). (c) $\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ and $\mathbf{U} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$ and indeed $\mathbf{X} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{B} = \begin{bmatrix} 4 \\ 14 \end{bmatrix}$.

2.26. (a) $(\mathbf{A}|\mathbf{B}) = \begin{bmatrix} 1 & 0 & 3 & 2 & 4 \\ -2 & 1 & 6 & 1 & 7 \\ -1 & 1 & 2 & -4 & 6 \\ 0 & 0 & -3 & 7 & 0 \end{bmatrix}$ and this can be reduced to the following echelon form: $\begin{bmatrix} 1 & 0 & 3 & 2 & 4 \\ 0 & 1 & 12 & 5 & 15 \\ 0 & 0 & 1 & 1 & 0.71429 \\ 0 & 0 & 0 & 1 & 0.21429 \end{bmatrix}$.
Backward substitution generates, in turn, $x_4 = 0.21429$, $x_3 = 0.5$, $x_2 = 7.92857$ and $x_1 = 2.07143$. Further reduction to reduced echelon form produces $\begin{bmatrix} 1 & 0 & 0 & 0 & 2.07143 \\ 0 & 1 & 0 & 0 & 7.92857 \\ 0 & 0 & 1 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0.21429 \end{bmatrix}$, with the solution $\mathbf{X}$ in the final column. For this problem,

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & 1 & 12 & 5 \\ 0 & 0 & -7 & -7 \\ 0 & 0 & 0 & 10 \end{bmatrix} \text{ and } \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0.4286 & 1 \end{bmatrix}$$

and it is easy to show that $\mathbf{X} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{B}$.

(b) Here $(\mathbf{A}|\mathbf{B}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 2 & 1 & 5 \\ 1 & 0 & 1 & 0 & 9 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$ and interchanging rows 2 and 3 makes things easier. After several elementary operations, you arrive at the following echelon and reduced echelon forms:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & -7 \\ 0 & 0 & 1 & 0.5 & 2.5 \\ 0 & 0 & 0 & 1 & -5 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & -5 \end{bmatrix}$$

and in either case it is easily established that $\mathbf{X} = \begin{bmatrix} 4 \\ -2 \\ 5 \\ -5 \end{bmatrix}$. For this example, $\mathbf{U} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}$, $\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0.5 & 1 \end{bmatrix}$ and, indeed, $\mathbf{X} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{B} = \begin{bmatrix} 4 \\ -2 \\ 5 \\ -5 \end{bmatrix}$.

2.27. The equations in Problem 1.12 are

$$\begin{aligned} -x_1 + x_2 &= 10 \\ -3x_1 + 2x_2 &= 16 \end{aligned}$$

If you try to use Gauss-Seidel on the equations *in this order* you will have

$$\begin{aligned} x_1 &= x_2 - 10 \\ x_2 &= 1.5x_1 + 8 \end{aligned}$$

and, for example, starting from $\mathbf{X}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, the iterations will explode toward negative infinity. However, if you write the original equations in the opposite order, then you will use Gauss-Seidel on the system

$$\begin{aligned} x_1 &= (2/3)x_2 - (16/3) \\ x_2 &= x_1 + 10 \end{aligned}$$

and the iterations converge nicely from $\mathbf{X}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to the (correct) solution $\mathbf{X} = \begin{bmatrix} 4 \\ 14 \end{bmatrix}$.

Chapter 3

3.1. (a) $f(x_0) = 4x_0^2 + 2x_0 + 1$; $f(x_0 + \Delta x) = 4(x_0 + \Delta x)^2 + 2(x_0 + \Delta x) + 1$. Therefore $\Delta y = f(x_0 + \Delta x) - f(x_0) = 4x_0^2 + 8x_0\Delta x + 4(\Delta x)^2 + 2x_0 + 2\Delta x + 1 - 4x_0^2 - 2x_0 - 1$ and so $\Delta y/\Delta x = 8x_0 + 4\Delta x + 2$. Taking the limit, as $\Delta x \to 0$, and dropping the subscript, since this is valid for any particular x_0, $dy/dx = 8x + 2$. Exactly the same approach will answer parts (b) and (c).

3.2. (a) $f'(x) = 6x + 12 = 0$, so $x^* = -2$. Since $f'(-3) = -6$ and $f'(1) = 6$, x^* constitutes a minimum point. (b) $f'(x) = 6x - 12 = 0$, so $x^* = 2$. Here $f'(1) = -6$ and $f'(3) = 6$, so x^* represents a minimum. (c) $f'(x) = -6x + 12 = 0$ leads to $x^* = 2$. Now $f'(1) = 6$ and $f'(3) = -6$, so x^* is a maximum point.

3.3. (a) $y - 430 = 72(x - 10)$, or $72x - y = 290$. (b) $y + 2 = 0$, or $y = -2$. (c) $y - 10 = -12(x - 4)$, or $12x + y = 58$.

3.4. (a) $f' = 6x^2 + 6x - 12 = 0$ leads to $x^* = -2$ or $x^{**} = 1$. $f'' = 12x + 6$, so $f''(x^*) < 0$ (x^* is a relative maximum) and $f''(x^{**}) > 0$ (x^{**} is a relative minimum). These extrema are relative, not absolute, because the function continues upward toward $+\infty$ to the right of x^{**} and downward toward $-\infty$ to the left of x^*.

(b) Setting $f'(x) = 3x^2(4x - 1) = 0$ yields $x^* = 0$ and $x^{**} = 0.25$. $f''(x^*) = 0$; taking higher order derivatives, $f'''(x^*) = -6$ and hence x^* represents a point of inflection. $f''(x^{**}) = 0$, so x^{**} represents a minimum. So $f(x)$ has no finite maximum but it has an absolute minimum at x^{**}.

(c) Setting $f'(x) = 3x^2 = 0$ yields $x^* = 0$; since $f''(x^*) = 0$, we find $f'''(x^*) = 6$, so x^* is a point of inflection.

(d) $f'(x) = 0$ leads to $x^* = -2$, where $f''(x^*) = 6$ and so x^* is an absolute minimum. The function extends upward infinitely far on both sides of x^*.

(e) $f'(x) = 0$ leads to $x^* = 0$ and $x^{**} = 0.75$. $f''(x^*) < 0$ (x^* is a maximum) and $f''(x^{**}) > 0$ (x^{**} is a minimum). Both are relative, not absolute, since the function extends downward to the left of x^* and upward to the right of x^{**}.

(f) $f'(x) = 4x^3 - 36x = 4x(x^2 - 9) = 0$, leading to $x^* =$

-3, $x^{**} = 0$ and $x^{***} = 3$. Here $f''(x) = 12x^2 - 36$, so $f''(x^*) > 0$, $f''(x^{**}) < 0$, and $f''(x^{***}) > 0$, meaning that x^* and x^{***} are minima and x^{**} is a maximum point. Since $f(x^*) = f(x^{***}) = -66$, these represent "absolute but equal" minima while $f(x^{**}) = 15$ represents a relative maximum. The function looks like a "rounded" W.

3.5. (a) Here $f'(x) = 0$ is a fifth degree equation with three real roots--$x^* = -1$, $x^{**} = 0.5$ and $x^{***} = 2$. $f''(x^*) = 0$ and $f'''(x^*) < 0$ (a point of inflection), $f''(x^{**}) > 0$ (a minimum point), and $f''(x^{***}) = 0$ and $f'''(x^{***}) > 0$ (another inflection point); x^{**} is a global minimum.

(b) Here $f'(x) = 0$ is a fourth degree equation, again with three real roots--$x^* = -3$, $x^{**} = 0$, and $x^{***} = 2$. Now $f''(x^*) = 0$ and $f'''(x^*) > 0$ (point of inflection), $f''(x^{**}) < 0$ (a maximum) and $f''(x^{***}) > 0$ (a minimum). Neither the maximum nor minimum point are absolute extrema.

3.6. (a) $x^* = 0.5$. (b) $x_1^* = x_2^* = 20$.

3.7. (a) $\mathbf{X}^* = \begin{bmatrix} -0.3333 \\ 0.6667 \end{bmatrix}$ is a minimum point. (b) $\mathbf{X}^* = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is a minimum point. (c) $\mathbf{X}^* = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ is neither a maximum nor a minimum. (d) $\mathbf{X}^* = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$ is a maximum point. (e) Here $\nabla f = [-6x_1^2 + 6x_2, 6x_1 - 2x_2 - 4]'$, so $\nabla f = 0$ is a set of nonlinear equations; the solutions are $\mathbf{X}^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{X}^{**} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$; $\mathbf{X}^*$ is neither a maximum nor a minimum; $\mathbf{X}^{**}$ is a maximum point.

3.8. From $z = xy/(x + y)$, $\partial z/\partial x = y^2/(x + y)^2$ and $\partial z/\partial y = x^2/(x + y)^2$, so $x(\partial z/\partial x) + y(\partial z/\partial y) = xy^2 + x^2y/(x + y)^2 = (xy)(y + x)/(x + y)^2 = xy/(x + y) = z$.

3.9. There should be 150 passengers for $R^* = \$1125$.

3.10. (a) $\mathbf{X}^* = \begin{bmatrix} 20 \\ 0 \end{bmatrix}$. (b) Since x_2 contributes only positively to the cost function that is to be minimized, it should remain as small as possible, namely at zero. (Negative total output generally has no meaning.) Examine the cost function at its two endpoints to x_1--0 and 12; $f(0, 0) = 500$ and $f(12, 0) =$

164, so set $x_1 = 12$ when there is a restriction that x_1 cannot exceed 12.

3.11. $\mathbf{X}^* = \begin{bmatrix} 26.67 \\ 6.67 \end{bmatrix}$.

3.12. (a) It was established in Problem 3.11 that there is only one stationary point, which is a minimum. So there is no finite maximum to the function. (b) If $x_1 = 12$, this becomes a function of one variable only, namely $f(x_2) = 4x_2^2 - 24x_2 + 164$. Setting $f'(x_2) = 0$ leads to $x_2^* = 3$, but this represents a minimum point. Clearly, for larger and larger values of x_2, $f(x_2)$ increases without limit. (c) Since $f(x_2)$ has a minimum at $x_2 = 3$, examine the value of the function at the relevant endpoints which lie on either side of $x_2 = 3$, namely at $x_2 = 0$ and at $x_2 = 5$. In the former case, $f(x_2) = 164$ and in the latter case, $f(x_2) = 144$, so it is best to set $x_2 = 0$.

3.13. Here $f'(x) = -4x + 90 = 0$ leads to $x^* = 22.5$--but $f''(x^*) = -4$, so x^* represents a maximum point. Therefore, examine $f(x) = -2x^2 + 90x + 600$ at the two endpoints, 0 and 50. Since $f(0) = 600$ and $f(50) = 100$, select $x = 50$. If the limits on x are $(0 \leq x \leq 40)$, and since $f(40) = 1000$, the answer is $x = 0$.

3.14. Here is one solution: $2x^3 - 3x^2 - 12x + 100$.

3.15. This is a function that should be drawn with a good computer graphics program. First-order conditions lead to $\mathbf{X}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and to $(x_1^{**})^2 + (x_2^{**})^2 = 1$, the set of values of x_1 and x_2 on a circle, centered at the origin, with radius 1. $\mathbf{X}^*$ indicates a minimum point and points on the circle are (relative) maxima.

3.16. $t^* \cong 32.7963$.

3.17. This problem should only be done if you have access to good computer software for finding derivatives and doing algebra.

Chapter 4

4.1. There is no finite maximum.

4.2. $\mathbf{X}^* = \begin{bmatrix} 3/31 \\ 18/31 \end{bmatrix}$.

4.3. $\mathbf{X}^* = \begin{bmatrix} 20/14 \\ 30/14 \\ 10/14 \end{bmatrix}$ represents a maximum.

4.4. $\mathbf{X}^* = \begin{bmatrix} 11 \\ 12 \end{bmatrix}$ is a minimum point.

4.5. $\mathbf{X}^* = \begin{bmatrix} 11/3 \\ 10/3 \\ 8 \end{bmatrix}$ is a minimum.

4.6. $\mathbf{X}^* = \begin{bmatrix} 230/150 \\ 334/150 \\ 38/150 \end{bmatrix} = \begin{bmatrix} 1.533 \\ 2.227 \\ 0.253 \end{bmatrix}$ is a maximum.

4.7. (a) $\mathbf{X}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. (b) $\mathbf{X}^* = \begin{bmatrix} 3/31 \\ 18/31 \end{bmatrix}$.

4.8. (a) $\mathbf{X}^* = \begin{bmatrix} 44 \\ 2 \end{bmatrix}$. (b) There is a local maximum at $\mathbf{X}^*$ from (a) but $f(\mathbf{X})$ has no finite maximum--for example, let $x_1 \to -\infty$ and $x_2 = 0$.

4.9. The solution is the same as for Problem 4.3.

4.10. The solution is the same as for Problem 4.6.

4.11. (a) There is no finite maximum; as $x \to -\infty$, the function grows ever larger and the constraint is met. (b) There is no finite maximum; as $x \to \infty$, $f(x)$ grows every larger and the constraint is met. (c) Same as (b). (d) At $x = -3$ there is a local minimum. (e) At $x = 3$ there is a local minimum. (f) The minimum occurs at the left-hand endpoint, namely at $x = 4$.

4.12. $\mathbf{X}^* = \begin{bmatrix} 25 \\ 25 \end{bmatrix}$.

4.13. (a) $\mathbf{X}^* = [10 \quad 10 \quad 10]'$. (b) $\mathbf{X}^* = [10 \quad 10 \quad 10]'$.

4.14. The problem is to minimize $2x_1x_2 + 2x_1x_3 + 2x_2x_3$ subject to $x_1x_2x_3 = c$, which leads to $x_1^* = x_2^* = x_3^* = c^{1/3}$.

4.15. $h^* = r^* = \lambda^* = c/(4 + \pi)$.

4.16. $x_1^* = x_2^* = 1.414k$.

Chapter 5

5.1. This is an exploratory problem that should be done on a computer, using either root finding software or a small program that you write.

5.2. (a) Newton-Raphson from $x^0 = 2$ finds $x^1 = 1.25$, $x^2 = 1.025$, $x^3 = 1.00030$ and $x^4 = 1.00000$ before stopping. (b) The secant method, from $x^0 = 2$ and $x^1 = 1.25$, finds $x^2 = 1.0769$, $x^3 = 1.0083$, $x^4 = 1.0003$ and $x^5 = 1.0000$. (c) Modified Newton takes 15 iterations. (d) Bisection requires a pair of initial points that bracket a root. Using $x^0 = 2$ and $x^1 = 0.5$, $x = 1.0000$ is reached in 18 iterations; if you use $x^0 = 2$ and $x^1 = -3$, $x = -1.0000$ is reached in 21 iterations.

5.3. (a) Newton-Raphson and Modified Newton fail because of the division by zero problem. The secant and bisection methods are unaffected; for bisection, the second of the two initial points must bracket a root.

(b) For Newton-Raphson, you end up at $x = -1$ in about 10 iterations. The secant method leads you to -1 or +1 depending on what you use for x^1, in conjunction with $x^0 = -0.01$. For example, with $x^1 = 2$, you get to $x = +1$, with $x^1 = -2$, you end up at $x = -1$. Modified Newton doesn't work; it explodes toward $x = +\infty$. Bisection, again, requires an initial bracketing pair. From $x^0 = -0.01$ and $x^1 = 2$, you reach $x = 1$ in about 17 iterations.

5.4. Since there is no real value for which $x^2 + 2 = 0$, all of the iterative methods simply fail to converge.

5.5. All of the iterative methods will find the solution at $x = 0$. (Modified Newton takes hundreds of iterations.)

5.6. For n = 2, the square of uncertainty with an area that is 10% of that of the unit square has sides of length 0.316 [which is $(0.1)^{(1/2)}$]. For n = 3, a volume that is 10% of that of the unit cube has sides of length 0.4642 [= $(0.1)^{(1/3)}$]. For n = 10 and n = 20, the corresponding "sides" have length 0.7943 and 0.8913, respectively [these are $(0.1)^{(1/10)}$ and $(0.1)^{(1/20)}$]. Remember that in each case, the length of the sides lies between 0 and 1, so by the time n = 20, the range of uncertainty for each x_i is almost 90% of its entire length.

5.7. The two solutions are, as indicated in the question, $\mathbf{X}^1 = \begin{bmatrix}-1.6794\\ 4.6794\end{bmatrix}$ and $\mathbf{X}^2 = \begin{bmatrix}2.6794\\ 0.3206\end{bmatrix}$, but of course the exact solution that you get will vary with the stopping criterion that you use.

5.8. Using a stopping criterion that the left hand sides of both equations should be within 10^{-5} of zero, the (only) solution is $\mathbf{X} = \begin{bmatrix}3.2424\\ 5.2297\end{bmatrix}$.

5.9. No solution.

5.10 (a) With the equations ordered (3, 1, 2, 4), the solutions for the unknowns are $x_1 = 9 - x_3$, $x_2 = 2 - x_1$, $x_3 = (5 - x_4)/2$, $x_4 = -x_3$. Convergence is fairly fast (on the order of 17 iterations) from a wide variety of starting points, and for the stopping criterion $\max_i [|f^i(\mathbf{X}^k)|] \leq 10^{-4}$. The only difference in the equation ordering (3, 1, 4, 2) is that now $x_3 = -x_4$ and $x_4 = 5 - 2x_3$, yet Gauss-Seidel now converges only from starting points that are very very close to the actual solution, $\mathbf{X} = [4\ -2\ 5\ -5]'$.

(b) Of the 8 possible equation orderings that are meaningful, only for (1, 3, 2, 4) does Gauss-Seidel converge, for a wide variety of starting points and the usual stopping criterion.

5.11. At $\mathbf{X}^k = \begin{bmatrix}1\\ 1\end{bmatrix}$, $\nabla f(\mathbf{X}^k) = \begin{bmatrix}8\\ 12\end{bmatrix}$.

(a) Forward difference approximations.

α	$f(x_1^k+\alpha,\ x_2)$	$f(x_1,\ x_2^k+\alpha)$	$f(\mathbf{X}^k)$	Approximation to ∇^k
0.2	-26.48	-25.64	-28	$\begin{bmatrix}7.6\\ 11.8\end{bmatrix}$
0.1	-27.22	-26.81	-28	$\begin{bmatrix}7.8\\ 11.9\end{bmatrix}$
0.01	-27.9202	-27.8801	-28	$\begin{bmatrix}7.98\\ 11.99\end{bmatrix}$
0.001	-27.9920	-27.9880	-28	$\begin{bmatrix}8\\ 12\end{bmatrix}$

(b) Backward difference approximations.

α	$f(x_1^k+\alpha,\ x_2)$	$f(x_1,\ x_2^k+\alpha)$	$f(\mathbf{X}^k)$	Approximation to ∇^k
0.2	-29.68	-30.44	-28	$\begin{bmatrix}8.4\\ 12.2\end{bmatrix}$
0.1	-28.82	-29.21	-28	$\begin{bmatrix}8.2\\ 12.1\end{bmatrix}$

0.01	-28.0802	-28.1201	-28	$\begin{bmatrix}8.02\\12.01\end{bmatrix}$
0.001	-28.0080	-28.0120	-28	$\begin{bmatrix}8\\12\end{bmatrix}$

(c) Central difference approximations. The central difference method, in this example, gets it right every time.

α	$[f(x_1^k+\alpha,\ x_2)-f(x_1^k-\alpha,\ x_2)]$	$[f(x_1,\ x_2^k+\alpha)-f(x_1,\ x_2^k-\alpha)]$	Approx. to ∇^k
0.2	3.2	4.8	$\begin{bmatrix}8\\12\end{bmatrix}$
0.1	1.6	2.4	$\begin{bmatrix}8\\12\end{bmatrix}$
0.01	0.16	0.24	$\begin{bmatrix}8\\12\end{bmatrix}$
0.001	0.016	0.024	$\begin{bmatrix}8\\12\end{bmatrix}$

5.12. For this problem, $\mathbf{H} = \mathbf{H}^k = \begin{bmatrix}-4 & 0\\0 & -2\end{bmatrix}$.

(a) The forward difference estimate of h_{11}^k is given by $[f(\mathbf{X}^k + 2\alpha\mathbf{I}_1) - 2f(\mathbf{X}^k + \alpha\mathbf{I}_1) + f(\mathbf{X}^k)]/\alpha^2$. If you work this out for α = 0.2, 0.1, 0.01 and 0.001, you will generate the exact value, -4, every time. (You need to carry twice as many places to the right of the decimal as there are in α; for example, for α = 0.001 you need six places to the right of the decimal point.) The forward difference estimate of h_{22}^k would be found in exactly the same way, with subscripts 1 replaced by 2. Backward difference estimates could be found similarly, subtracting rather than adding α.

The forward difference estimate of h_{12}^k is given by $[f(\mathbf{X}^k + \alpha\mathbf{I}_1 + \alpha\mathbf{I}_2) - f(\mathbf{X}^k + \alpha\mathbf{I}_1) - f(\mathbf{X}^k + \alpha\mathbf{I}_2) + f(\mathbf{X}^k)]/\alpha^2$. Again, for α = 0.2, 0.1, 0.01 and 0.001 you will get exactly 0, which is h_{12}. The forward difference estimate for h_{21} would also be 0 for all values of α. Again, backward difference estimates result if α is subtracted rather than added in all steps.

(b) If you work out the details of the central difference method, you will also find that the estimate is exactly right for all elements of $\mathbf{H}$. When $\mathbf{H}$ is made up of constants, as here, the estimates provided by forward difference methods and central difference methods (and backward difference methods, also), are identical and exactly correct.

5.13. (a) At $\mathbf{X}^k = \begin{bmatrix}1\\1\end{bmatrix}$, $\nabla f(\mathbf{X}^k) = \begin{bmatrix}-4\\13\end{bmatrix}$.

(i) Forward difference approximations.

α	$f(x_1^k+\alpha, x_2)$	$f(x_1, x_2^k+\alpha)$	$f(\mathbf{X}^k)$	Approximation to ∇^k
0.2	2.088	6.184	3	$\begin{bmatrix}-4.56\\15.92\end{bmatrix}$
0.01	2.9597	3.131403	3	$\begin{bmatrix}-4.03\\13.14\end{bmatrix}$
0.001	2.995997	3.013014	3	$\begin{bmatrix}-4.003\\13.014\end{bmatrix}$

(ii) Backward difference approximations.

α	$f(x_1^k-\alpha, x_2)$	$f(x_1, x_2^k-\alpha)$	$f(\mathbf{X}^k)$	Approximation to ∇^k
0.2	3.672	0.936	3	$\begin{bmatrix}-3.36\\10.32\end{bmatrix}$
0.01	3.039699	2.871397	3	$\begin{bmatrix}-3.97\\12.86\end{bmatrix}$
0.001	3.003997	2.987014	3	$\begin{bmatrix}-3.997\\12.986\end{bmatrix}$

(iii) Central difference approximations.

α	$[f(x_1^k+\alpha, x_2)-f(x_1^k-\alpha, x_2)]$	$[f(x_1, x_2^k+\alpha)-f(x_1, x_2^k-\alpha)]$	Approx. to ∇^k
0.2	-1.584	5.248	$\begin{bmatrix}-3.96\\13.12\end{bmatrix}$
0.01	-0.079999	0.260006	$\begin{bmatrix}-4\\13\end{bmatrix}$
0.001	-0.007999	0.026000	$\begin{bmatrix}-4\\13\end{bmatrix}$

(b) At $\mathbf{X}^k = \begin{bmatrix}1\\1\end{bmatrix}$, $\mathbf{H}^k = \begin{bmatrix}-6 & -2\\-2 & 28\end{bmatrix}$. As in Problem 5.11, the forward difference estimate of h_{11}^k is given by $[f(\mathbf{X}^k + 2\alpha I_1) - 2f(\mathbf{X}^k + \alpha I_1) + f(\mathbf{X}^k)]/\alpha^2$ and the estimate of h_{12}^k is given by $[f(\mathbf{X}^k + \alpha I_1 + \alpha I_2) - f(\mathbf{X}^k + \alpha I_1) - f(\mathbf{X}^k + \alpha I_2) + f(\mathbf{X}^k)]/\alpha^2$. There are similar expressions for h_{21}^k and h_{22}^k. Backward and central difference estimates can also easily be found.

For h_{11}^{k},

α	$f(\mathbf{X}^k + 2\alpha I_1)$	$2f(\mathbf{X}^k + \alpha I_1)$	$f(\mathbf{X}^k)$	Approx. to h_{11}^{k}
0.2	0.984	4.176	3	-4.8
0.01	2.9188	5.9194	3	-6
0.001	2.991988	5.991994	3	-6

For h_{12}^{k},

α	$f(\mathbf{X}^k + \alpha I_1 + \alpha I_2)$	$f(\mathbf{X}^k + \alpha I_1)$	$f(\mathbf{X}^k + \alpha I_2)$	$f(\mathbf{X}^k)$	Approx. to h_{12}^{k}
0.2	5.184	2.088	6.184	3	-2.2
0.01	3.090903	2.9597	3.131403	3	-2
0.001	3.009009	2.995997	3.013014	3	-2

Results would be similar for h_{21}^{k} and h_{22}^{k}, and, again, backward and central difference estimates could also be found..

5.14. Straightforward substitution of $f'(x^k) \cong [f(x^k + \alpha) - f(x^k)]/\alpha$ and $f'(x^k + \beta) \cong [f(x^k + \beta + \alpha) - f(x^k + \beta)]/\alpha$ into $f''(x^k) \cong [f'(x^k + \beta) - f'(x^k)]/\beta$ yields the result. When $\alpha = 0.1$, $\beta = 0.2$ and $x^k = 2$, the estimate is $(3.29 - 2.84 - 2.41 + 2)/0.02 = 2$, which is exactly right; in this example $f''(x) = 2$ everywhere.

Chapter 6

6.1. $I^0 = (0, x_n)$, with midpoint x_1. $I^1 = (0, x_1 + (\delta/2)) = (0.5)I^0 + (\delta/2)$. $I^2 = (0.5)I^1 + (\delta/2)$, and substituting I^1 gives $I^2 = (0.5)^2 I^0 + (0.5)(\delta/2) + (\delta/2)$. The results for $I^3, I^4, \ldots, I^k$ follow directly.

6.2. This is an exploratory question. Continuing the simplex geometry will take you ever closer to the minimum at $\mathbf{X}^* = \begin{bmatrix} 2 \\ 2.5 \end{bmatrix}$.

6.3. In all cases, $x = 0$ represents a local maximum and $x = \pm 3$ are local minima.

6.4. $\mathbf{X}^* = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$ is the maximum point. A simple gradient approach--$\mathbf{X}^{k+1} = \mathbf{X}^k + (0.25)\nabla^k$--from $\mathbf{X}^0 = 0$, leads to the solution in 21 iterations, using a stopping criterion that both elements of $\nabla^k \le 10^{-5}$.

6.5. (a) $\mathbf{X}^* = \begin{bmatrix} -0.3333 \\ 0.6667 \end{bmatrix}$ is the minimum point. Using the gradient approach in (6-1) leads to this solution in 17 iterations, from $\mathbf{X}^0 = 0$ and with a stopping criterion that both elements in $\nabla^k \le 10^{-5}$.

(b) $\mathbf{X}^* = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is the minimum. Using (6-1), from $\mathbf{X}^0 = 0$ and with the same stopping criterion as in (a), leads to $\mathbf{X}^*$ in 24 iterations.

If you are using a program which easily allows you to change α in both parts of this problem, you might do so and observe how sensitive the iterative procedure is to the choice of α. That is why most gradient-like methods use a step size (α) that varies with each step.

6.6. (a) For a maximization problem for a function of one variable, (6-1) becomes $x^{k+1} = x^k + \Delta x^k = x^k + \alpha^k f'(x^k)$, where $f'(x^k)$ plays the role of ∇^k in the multivariable case in (6-1) and the iterative scheme works out to be

$$x^{k+1} = x^k + [-1/f''(x^k)]f'(x^k)$$

For a minimization problem for $f(x)$, the steps are

$$x^{k+1} = x^k + [1/f''(x^k)][-f'(x^k)]$$

(b) The iterative steps are $x^{k+1} = x^k + (1/6)(-6x^k + 6)$.

You will get to $x^* = 1$ in one iteration from any of the starting points, since the quadratic approximation given by a second-order Taylor Series is exact for a quadratic $f(x)$ such as here.

(c) This is just the "negative" of the problem in (b). Since this is a minimization problem, movement follows $x^{k+1} = x^k + \alpha^k[-f'(x^k)]$ and $\alpha^k = -1/6$, so $x^{k+1} = x^k + (1/6)(6x^k - 6)$ and, as in (c), you will get to $x^* = 1$ in one iteration from any of the starting points.

6.7. The parallel to Newton's method in (6-7), for a function of one variable, is

$$x^{k+1} = x^k + [-f'(x^k)/f''(x^k)]$$

for both a maximization problem and a minimization problem. Cauchy and Newton are the same for the $f(x)$ case.

6.8. (a) From $\mathbf{X}^{k+1} = \mathbf{X}^k + \mathbf{C}^k$, using a second-order Taylor Series approximation,

$$\Delta f(\mathbf{X}^k) = f(\mathbf{X}^{k+1}) - f(\mathbf{X}^k) = [\nabla^k]'[\mathbf{C}^k] + (0.5)[\mathbf{C}^k]'\mathbf{H}^k[\mathbf{C}^k]$$

so

$$\partial\Delta f(\mathbf{X}^k)/\partial\mathbf{C}^k = \nabla^k + \mathbf{H}^k\mathbf{C}^k = 0$$

leads to

$$\mathbf{C}^k = -(\mathbf{H}^k)^{-1}[\nabla^k]$$

which is exactly Newton.

Second-order conditions depend on $\partial^2\Delta f(\mathbf{X}^k)/(\partial\mathbf{C}^k)^2 = \mathbf{H}^k$, so we need $\mathbf{H}^k$ negative definite for a maximum and positive definite for a minimum.

(b) Now $\mathbf{X}^{k+1} - \mathbf{X}^k = \bar{\alpha}\Delta\mathbf{X}^k$, so the Taylor Series approximation is

$$\Delta f(\mathbf{X}^{k+1}) = f(\mathbf{X}^{k+1}) - f(\mathbf{X}^k)$$
$$= [\nabla^k]'[\bar{\alpha}\Delta\mathbf{X}^k] + (0.5)[\bar{\alpha}\Delta\mathbf{X}^k]'\mathbf{H}^k[\bar{\alpha}\Delta\mathbf{X}^k]$$

From

$$d\Delta f(\mathbf{X}^k)/d\Delta\mathbf{X}^k = \bar{\alpha}\nabla^k + \bar{\alpha}^2\mathbf{H}^k[\Delta\mathbf{X}^k] = 0$$

we have

$$\Delta\mathbf{X}^k = -(1/\bar{\alpha})(\mathbf{H}^k)^{-1}[\nabla^k]$$

Second-order conditions come from $d^2\Delta f(\mathbf{X}^k)/(d\Delta\mathbf{X}^k)^2 = (\bar{\alpha}^k)^2\mathbf{H}^k$, so, again $\mathbf{H}^k$ must be positive definite for a minimum and negative definite for a maximum.

6.9. (a) Straightforward substitution of $\mathbf{x}^k$ and $\mathbf{g}^k$ into (6-32) leads to $\mathbf{B}^{k+1}$ as shown. This is a positive definite matrix; all leading principal minors are positive.

(b) Substitution into (6-32) again leads to the $\mathbf{B}^{k+1}$ shown in the problem. While the first two leading principal minors are positive, $|\mathbf{B}^{k+1}| = -1$, so the matrix is not positive definite.

6.10. If you begin by letting $(\mathbf{g}^k)'\mathbf{B}^k\mathbf{g}^k = s$ (a scalar) and $(\mathbf{x}^k)'\mathbf{g}^k = t$ (another scalar), the algebra is simpler. Then expand the two expressions, in (6-41) and (6-42), removing parentheses and carrying out the multiplications, and remember that $\mathbf{B}^k$ is symmetric, so $\mathbf{B}^k = (\mathbf{B}^k)'$. The first two terms in both expressions are $\mathbf{B}^k + \mathbf{x}^k(\mathbf{x}^k)'/t$. The three remaining terms in (6-41) are matched by terms in (6-42), in a different order. In addition, (6-42) contains the term $[\mathbf{B}^k\mathbf{g}^k(\mathbf{g}^k)'\mathbf{B}^k/s]$, both added and subtracted.

Chapter 7

7.1. $\mathbf{X}^* = [0 \quad 0 \quad 3.6]'$, $\mathbf{S}^* = [8.4 \quad 0]'$ and $f(\mathbf{X}^*) = 54$.

7.2. $\mathbf{X}^* = \begin{bmatrix} 0 \\ 12/7 \end{bmatrix}$, $\mathbf{S}^* = \begin{bmatrix} 0 \\ 5/7 \end{bmatrix}$ and $f(\mathbf{X}^*) = 13\ 5/7$.

7.3. $\mathbf{X}^* = \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}$, $\mathbf{S}^* = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ and $f(\mathbf{X}^*) = 13.5$.

7.4. $\mathbf{X}^* = [0 \quad 5 \quad 10 \quad 0]'$, $\mathbf{S}^* = \mathbf{0}$ and $f(\mathbf{X}^*) = 140$.

7.5. The feasible region for this problem is unbounded; there is no finite maximum.

7.6. This is an exploratory problem. The optimal solutions to all parts of the problem should appear after three pivots.

7.7. (a) Minimize $12y_1 + 18y_2$ subject to $y_1 + 2y_2 \geq 3$, $y_1 + y_2 \geq 2$, $y_1 + 5y_2 \geq 15$ and $\mathbf{Y} \geq \mathbf{0}$. $\mathbf{Y}^* = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, $\mathbf{T}^* = [3 \quad 1 \quad 0]'$ and $\Delta^* = 54$. Recalling that $\mathbf{X}^* = [0 \quad 0 \quad 3.6]'$ and $\mathbf{S}^* = [8.4 \quad 0]'$, it is obvious by inspection that $\mathbf{Y}^* \circ \mathbf{S}^* = 0$ and $\mathbf{X}^* \circ \mathbf{T}^* = 0$.

(b) Maximize $12y_1 + y_2$ subject to $2y_1 + 2y_2 \leq 6$, $7y_1 + y_2 \leq 8$ and $\mathbf{Y} \geq 0$. $\mathbf{Y}^* = \begin{bmatrix} 8/7 \\ 0 \end{bmatrix}$, $\mathbf{T}^* = \begin{bmatrix} 26/7 \\ 0 \end{bmatrix}$ and $\Delta^* = 13\ 5/7$. Since $\mathbf{X}^* = \begin{bmatrix} 0 \\ 12/7 \end{bmatrix}$, and $\mathbf{S}^* = \begin{bmatrix} 0 \\ 5/7 \end{bmatrix}$, $\mathbf{Y}^* \circ \mathbf{S}^* = 0$ and $\mathbf{X}^* \circ \mathbf{T}^* = 0$.

(c) Maximize $2y_1 + 6y_2$ subject to $3y_1 + y_2 \leq 5$, $4y_1 + 4y_2 \leq 9$ and $\mathbf{Y} \geq 0$. $\mathbf{Y}^* = \begin{bmatrix} 0 \\ 2.25 \end{bmatrix}$, $\mathbf{T}^* = \begin{bmatrix} 2.75 \\ 0 \end{bmatrix}$ and $\Delta^* = 13.5$. Since $\mathbf{X}^* = \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}$ and $\mathbf{S}^* = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\mathbf{Y}^* \circ \mathbf{S}^* = 0$ and $\mathbf{X}^* \circ \mathbf{T}^* = 0$.

(d) Minimize $10y_1 + 30y_2$ subject to $2y_1 \geq 1$, $2y_2 \geq 8$, $y_1 + 2y_2 \geq 10$, $5y_2 \geq 9$ and $\mathbf{Y} \geq \mathbf{0}$. $\mathbf{Y}^* = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\mathbf{T}^* = [3 \quad 0 \quad 0 \quad 11]'$; since $\mathbf{X}^* = [0 \quad 5 \quad 10 \quad 0]'$, $\mathbf{S}^* = 0$, $\mathbf{Y}^* \circ \mathbf{S}^* = 0$ and $\mathbf{X}^* \circ \mathbf{T}^* = 0$.

(e) Minimize $6y_1 + 16y_2$ subject to $-y_1 - y_2 \geq 1$, $y_1 + 2y_2 \geq 2$ and $\mathbf{Y} \geq \mathbf{0}$. The feasible region for this problem is empty. This follows from the fact that the primal is unbounded.

7.8. The solution space geometry for the two primal problems will be the same, so the optimal solution to the primal is unchanged. Therefore the optimal value of the dual objective function is the same for both problems, and the same dual variables will be positive and zero in both dual problems. The primal problem whose i-th constraint is k times the i-th constraint of the other will therefore have an associated dual variable that is (1/k)-th as large, provided it is nonzero in the optimal solution. (If it is zero, it will remain zero.) You can easily verify this for a numerical example.

7.9. (a) The solution space geometry for the primal will be unchanged and so the optimal values of the primal variables would not be affected at all. (b) In the solution-space picture for the dual, the right-hand sides of all constraints would be (1/10)-th as large; hence optimal values of the dual variables would also be (1/10)-th as large. You can easily verify this also in the same numerical example.

7.10. This is a derivation question. Given the initial pivot column (ℓ) and row (k), it requires that you go through the algebra of first expressing x_ℓ as a linear function of all the other x's along with s_k. This generates the first two rules in Table 7.6. Using this result for x_ℓ in any other row of the original simplex table will produce the last two results in Table 7.6.

7.11. Consider a small general problem.

$$\begin{aligned} \text{Maximize} \quad & p_1x_1 + p_2x_2 \\ \text{subject to} \quad & a_{11}x_1 + a_{12}x_2 \le b_1 \\ & a_{21}x_1 + a_{22}x_2 \le b_2 \end{aligned}$$

and $\mathbf{X} \ge 0$. Converting to augmented form

$$\begin{aligned} \text{Maximize} \quad & p_1x_1 + p_2x_2 + 0s_1 + 0s_2 \\ & a_{11}x_1 + a_{12}x_2 + s_1 = b_1 \\ & a_{21}x_1 + a_{22}x_2 + s_2 = b_2 \end{aligned}$$

The dual constraints associated with the coefficient columns for the slack variables will be

$$1y_1 + 0y_2 \ge 0;\ \text{that is } y_1 \ge 0$$
$$0y_1 + 1y_2 \ge 0;\ \text{that is } y_2 \ge 0$$

7.12. By simply observing that the first dual constraint is $-5y_1 - y_2 \ge 2$, which is clearly impossible for $\mathbf{Y} \ge 0$, you can see that the dual problem has no feasible solution. If you try to solve the primal problem, you will find the same signal as in Problem 7.5 that the primal feasible region is unbounded.

7.13. 140 Clippers, 150 Cruisers, $\Pi^* = 430{,}000$.

7.14. 100 Clippers, 200 Cruisers, $\Pi^* = 300{,}000$.

7.15. (a) 40 in A, 18 in B, $\Pi^* = 24{,}600$. (b) $y_1^* = 7$; the range over which this marginal figure is valid is $1200 \le b_1 \le 3200$. Therefore an additional 40 person-hours would be worth \$280. (c) Since $y_2^* = 4.5$ and $y_3^* = 0$, it should be made available in region A.

7.16. (a) No. An optimal solution is a basic feasible solution, and so at most two of the three unknowns will be nonzero.

(b) $\mathbf{X}^* = [30 \quad 15 \quad 0]'$, $\Pi^* = 24{,}000$.

(c) There is no conflict. The optimal solution in (b) uses 810 pounds of recycled plastic.

(d) Decreased by 555.56, the element under x_3 in the top row of the optimal simplex table.

(e) In the optimal solution, $y_2^* = 33.33$, with a range $180 \le b_2 \le 450$. So the strike would cost 333.33 [= $(360 - 350)y_2^*$]. If b_2 went down to 150, that is below the range of validity for y_2^*.

(f) The range over which changes in the value of p_3 will not alter the current optimal solution has an upper limit of 755.56, so increases to 500 or to 750 would not change $\mathbf{X}^*$.

7.17. (a) The problem is to maximize $x_{AJ} + x_{AK}$, subject to upper limit constraints on each of the segments and also on constraints assuring that everything that flows into a pumping station must flow out from it and subject also to nonnegativity of all the flows. The upper limits are:

$$x_{AJ} \le 20,\ x_{AK} \le 12,\ x_{JL} \le 16,\ x_{LB} \le 12,$$
$$x_{KM} \le 8,\ x_{MB} \le 32,\ x_{JM} \le 24,\ x_{KL} \le 28$$

The inflow = outflow constraints are:

$$x_{AJ} = x_{JL} + x_{JM},\ x_{AK} = x_{KL} + x_{KM},\ x_{JL} + x_{KL} = x_{LB},\ x_{KM} + x_{JM} = x_{MB}$$

There are multiple optima to this problem. One optimal solution is

$$x_{AJ}^* = 20,\ x_{AK}^* = 12,\ x_{JL}^* = 8,\ x_{LB}^* = 12,$$
$$x_{KM}^* = 8,\ x_{MB}^* = 20,\ x_{JM}^* = 12,\ x_{KL}^* = 4$$

so the maximum amount that can be shipped is 32 barrels.

(b) Since the two pipelines out of A are used to capacity in the optimal solution, the increase in capacity would be needed there. From the optimal solution, the dual variables on the lines from A to J and from A to K are both 1; the right-hand side ranges are 8 to 32 for the capacity on AJ and 8 to 20 for the capacity on AK. So you could put all 5 units of new capacity on either AJ or AK. In either case, total flow would increase by 5, to 37.

(c) The problem now is to minimize

$$20x_{AJ} + 12x_{AK} + 16x_{JL} + 12x_{LB} + 8x_{KM} + 32x_{MB} + 24x_{JM} + 28x_{KL}$$

subject to the same inflow = outflow constraints at each pumping station along with an upper limit of one on the flow along each line and the restriction that either x_{AJ} or x_{AK} must be 1. The new capacity constraints are simply

$$x_{AJ} \le 1,\ x_{AK} \le 1,\ x_{JL} \le 1,\ x_{LB} \le 1,$$
$$x_{KM} \le 1,\ x_{MB} \le 1,\ x_{JM} \le 1,\ x_{KL} \le 1$$

To restrict the unit flow to either AJ or AK, add the constraint $x_{AJ} + x_{AK} = 1$. The optimal solution is $x^*_{AJ} = x^*_{JL} = x^*_{LB} = 1$, for a total cost of 48.

Chapter 8

8.1. The optimal *basic* solutions are $\mathbf{X}^* = \begin{bmatrix}2\\6\end{bmatrix}$ and $\mathbf{X}^{**} = \begin{bmatrix}8\\2\end{bmatrix}$; in both cases, $\Pi^* = 66$. All *nonbasic* solutions on the line connecting $\mathbf{X}^*$ and $\mathbf{X}^{**}$ are also equally optimal. (Geometrically, the objective function in this problem is parallel to the third constraint.)

8.2. The optimal solution to the primal problem is degenerate; $\mathbf{X}^* = [6 \quad 3 \quad 0]'$. The zero in the optimal solution indicates that there was more than one "smallest" ratio in the previous pivot column. Choosing the first as the pivot element, $\mathbf{X}^*$ results, along with $\mathbf{Y}^* = [10/7 \quad 9/7 \quad 0]'$; choosing the second as the pivot, the same $\mathbf{X}^*$ results, along with $\mathbf{Y}^* = [0 \quad 1/3 \quad 10/3]'$.

8.3. $\mathbf{X}^* = \begin{bmatrix}6.872\\7.436\end{bmatrix}$, $\Pi^* = 21.179$.

8.4. $\mathbf{X}^* = \begin{bmatrix}5\\2\end{bmatrix}$, $\Pi^* = 64$.

8.5. $\mathbf{X}^* = [0 \quad 3 \quad 3]'$, $\Pi^* = 51$.

8.6. The problem is to

Maximize $100x_A + 130x_B + 120x_C + 125x_D + 110x_E + 135x_F$

subject to

$$15000x_A + 15000x_B + 12000x_C + 18000x_D + 19000x_E + 17000x_F \le 40{,}000$$

and, for noncontiguity,

$$x_A + x_B \le 1,\ x_C + x_D \le 1,$$
$$x_C + x_E \le 1,\ x_C + x_F \le 1,\ x_D + x_F \le 1$$

and, for all x, $0 \le x \le 1$ and integer. The optimal solution is $x_B^* = x_F^* = 1$ and all other variables are zero; the optimal value of the objective function is 265.

8.7. The problem can be formulated as

$$\begin{aligned} \text{Maximize} \quad & 6000x_A + 5000x_B \\ \text{subject to} \quad & 2000x_A + 3000x_B \le 4000 \\ & x_A \le 1 \\ & x_B \le 1 \end{aligned}$$

and $\mathbf{X} \ge 0$ and integer

Using either Gomory or branch and bound, you will find $\mathbf{X}^* = \begin{bmatrix}1\\0\end{bmatrix}$.

8.8. Here $x^*_{11} = 15$, $x^*_{12} = 0$, $x^*_{13} = 15$, $x^*_{21} = 5$, $x^*_{22} = 45$, $x^*_{23} = 0$; minimal costs are 855.

8.9. This problem has multiple optima. One solution is $x^*_{12} = 4$, $x^*_{14} = 10$, $x^*_{16} = 4$, $x^*_{21} = 8$, $x^*_{22} = 2$, $x^*_{23} = 6$, $x^*_{25} = 7$, $x^*_{36} = 6$, and all other variables are zero; minimal costs are 54.

8.10. Optimal assignments are Truck 1 to D, Truck 2 to C, Truck 3 to B and Truck 4 to A; minimized costs are 4300.

8.11. The optimal assignment is W_1 to L_5, W_2 to L_3, W_3 gets no assignment, W_4 to L_4, W_5 to L_1, W_r to L_2, for a total cost of 210.

8.12. (a) $v^*_1 = -3$. (b) $u^*_1 = +1$.

8.13. $x^*_{12} = 6$, $x^*_{14} = 8$, $x^*_{16} = 4$, $x^*_{21} = 8$, $x^*_{23} = 6$, $x^*_{24} = 2$, $x^*_{25} = 7$, $x^*_{36} = 6$; minimized cost is 60.

8.14. Using OW^*_{ij} to indicate the optimal shipment from origin i to warehouse j and WD^*_{jk} for the optimal shipment from warehouse j to destination k, the optimal solution is $x^*_{12} = 6$, $x^*_{14} = 10$, $x^*_{15} = 2$, $x^*_{23} = 6$, $x^*_{36} = 6$, $OW^*_{22} = 17$, $WD^*_{21} = 8$, $WD^*_{25} = 5$, $WD^*_{26} = 4$; total cost is now 36.75.

8.15. (a) The optimal shipping matrix is

		Region 1			Region 2	
		1	2	3	1	2
Region 1	1			100		
	2	100		100		
Region 2	1				150	
	2				100	50
	3		100			

and minimized total shipping costs = 29,000.

(b) The optimal shipments matrix is the same; costs are now 39,000.

(c) The optimal shipments matrix is now

		Region 1			Region 2	
		1	2	3	1	2
Region 1	1			100		
	2	100		100		
Region 2	1		50		50	50
	2				150	
	3		50		50	

with a total cost of 30,000.

(d) The optimal shipments matrix is now

		Region 1			Region 2	
		1	2	3	1	2
Region 1	1			100		
	2	100		100		
Region 2	1		100			50
	2				150	
	3				100	

with total shipping costs of 30,500.

(e) Region 1, City 2; the increase would be 100.

(f) Region 1, Plant 1; the decrease would be 70.

8.16. (a) Plants 1 and 3 are both cheapest at location B; Locations A and C are both cheapest for plant 2. (b) The optimal assignment is: 1 to B, 2 to C and 3 to A; total cost = 12. (c) In this case, 3 should also go to B; total cost = 10. (d) Now the optimal assignment is 1 to D, 2 to C, 3 to B; total cost = 9.

8.17. This is an exploratory question. Your results will depend on the additional objective functions that you formulate and the trade-off coefficients that you use.

8.18. This also is an exploratory problem.

8.19. This also is an exploratory problem.

Chapter 9

9.1. After seven pivots, $\mathbf{X}^* = [10{,}000 \quad 0 \quad 0]'$, $\Pi^* = 10{,}000$. Notice that if you used, as the criterion for pivot column selection, the *greatest total increase* in the objective function (not the greatest *unit* increase, which is what is used in the text), you would choose column 1 for the initial pivot, and that would lead directly to the optimal solution.

9.2. The problem is

$$\begin{aligned} \text{Maximize} \quad & x_1 + 10x_2 + 100x_3 + 1000x_4 \\ \text{subject to} \quad & x_1 + 20x_2 + 200x_3 + 2000x_4 \le 10^6 \\ & x_2 + 20x_3 + 200x_4 \le 10^4 \\ & x_3 + 20x_4 \le 10^2 \\ & x_4 \le 10^0 \\ & \text{and } \mathbf{X} \ge \mathbf{0} \end{aligned}$$

The procedure is identical to that in Problem 9.1, only now it will take 15 rather than seven pivots.

9.3. (a) $\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & d_{nn} \end{bmatrix}$, where $d_{jj} > 0$. So $\mathbf{D}$ is positive definite if and only if all leading principal minors are positive [this is (1-40) in Chapter 1]. Here $|\mathbf{D}_1| = d_{11} > 0$; $|\mathbf{D}_2| = d_{11}d_{22} > 0$; $|\mathbf{D}_3| = d_{11}d_{22}d_{33} > 0$. Continuing, using the fourth row along which to evaluate $|\mathbf{D}_4|$, $|\mathbf{D}_4| = 0 + 0 + 0 + d_{44}|\mathbf{D}_3| > 0$, $\dots$, $|\mathbf{D}_n| = 0 + 0 + \dots + 0 + d_{nn}|\mathbf{D}_{n-1}| > 0$.

(b) Here $\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$, with $m < n$ and $\rho(\mathbf{A}) = m$.

Using $\mathbf{A}_i$ for the i-th column of $\mathbf{A}$ and ${}_j\mathbf{A}$ for the j-th row of $\mathbf{A}$,

$$\mathbf{AA}' = \begin{bmatrix} ({}_1\mathbf{A})(\mathbf{A}')_1 & ({}_1\mathbf{A})(\mathbf{A}')_2 & \dots & ({}_1\mathbf{A})(\mathbf{A}')_m \\ ({}_2\mathbf{A})(\mathbf{A}')_1 & ({}_2\mathbf{A})(\mathbf{A}')_2 & \dots & ({}_2\mathbf{A})(\mathbf{A}')_m \\ \vdots & \vdots & & \vdots \\ ({}_m\mathbf{A})(\mathbf{A}')_1 & ({}_m\mathbf{A})(\mathbf{A}')_2 & \dots & ({}_m\mathbf{A})(\mathbf{A}')_m \end{bmatrix}$$

(i) Notice that $({}_i\mathbf{A}) = [a_{i1}, \ldots, a_{in}]$, $(\mathbf{A}')_j = \begin{bmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{bmatrix}$, $({}_i\mathbf{A})(\mathbf{A}')_j =$ $(a_{i1}a_{j1} + \ldots + a_{in}a_{jn})$; also $({}_j\mathbf{A}) = [a_{j1}, \ldots, a_{jn}]$, $(\mathbf{A}')_i = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{in} \end{bmatrix}$, and $({}_j\mathbf{A})(\mathbf{A}')_i = (a_{j1}a_{i1} + \ldots + a_{jn}a_{in}) = ({}_i\mathbf{A})(\mathbf{A}')_j$, so $\mathbf{AA}'$ is symmetric. (ii) If $\rho(\mathbf{A}) = m$, the rows of $\mathbf{A}$ are linearly independent. Suppose the opposite, that ${}_1\mathbf{A} = k[{}_2\mathbf{A}]$. Then, for example, the second leading principal minor of $\mathbf{AA}' =$ $[k({}_2\mathbf{A})k(\mathbf{A}')_2][({}_2\mathbf{A})(\mathbf{A}')_2] - [k({}_2\mathbf{A})(\mathbf{A}')_2]^2$, which is easily shown to be equal to zero, violating the requirement that all principal minors be positive for a positive definite matrix.

$$\text{(c)}\quad \mathbf{ADA}' = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}.$$

The argument here is the same as in (b), except that now there is also a set of d_{jj} terms involved. But since all $d_{jj} > 0$, the results are not changed.

(d) Here is an example in which $m = 3$, $n = 5$ and $\rho(\mathbf{A}) = 2 < m$: $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 8 & 7 & 6 & 5 & 4 \end{bmatrix}$. Here $\mathbf{AA}' = \begin{bmatrix} 55 & 110 & 80 \\ 110 & 220 & 160 \\ 80 & 160 & 190 \end{bmatrix}$, which is symmetric but also singular; $\rho(\mathbf{AA}') = 2$ (for example, row 2 is twice row 1).

9.4. $a^* = b^* = c^* = (2/3)\bar{s}$.

9.5. Case 1. $R = 2.3094$ and $r = 1.1547$ so $R/r = 2$. $A_c = \pi r^2 = 4.1888$, $A_t = 6.9282$, so $A_c/A_t = 0.6046$. Case 2. $R = 2.5515$ and $r = 0.8165$ so $R/r = 3.1250$. $A_c = 2.0944$, $A_t = 4.8990$, so $A_c/A_t = 0.4275$. Case 3 $R = 3.0505$ and $r = 0.7939$ so $R/r = 3.8426$. $A_c = 1.9801$, $A_t = 4.7634$, so $A_c/A_t = 0.4157$.

9.6. (a) $\mathbf{A}_R = (\mathbf{A}')(\mathbf{AA}')^{-1} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \{[1\ 2\ 1] \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\}^{-1} = \begin{bmatrix} 1/6 \\ 2/6 \\ 1/6 \end{bmatrix}$, so $\tilde{\mathbf{X}}^0 = (\mathbf{A}')(\mathbf{AA}')^{-1}\mathbf{B} = \begin{bmatrix} 10/6 \\ 20/6 \\ 10/6 \end{bmatrix}$. The distance of this $\tilde{\mathbf{X}}^0$ to the origin is $[(10/6)^2 + (20/6)^2 + (10/6)^2]^{0.5} = 4.08$. From each of the endpoints of the constraint, $\tilde{\mathbf{X}}^0 = \begin{bmatrix}10\\0\\0\end{bmatrix}, \begin{bmatrix}0\\5\\0\end{bmatrix}$ or $\begin{bmatrix}0\\0\\10\end{bmatrix}$, the distances to the origin are simply 10, 5 and 10, respectively. From a randomly chosen alternative point on the constraint, $\tilde{\mathbf{X}}^0 = \begin{bmatrix}3\\1\\5\end{bmatrix}$, the distance to the origin is 5.92. You can try alternative points on the constraint; they will all be more than 4.08 units away from the origin.

(b) The constraint $x_1 + x_2 + x_3 = 12$ is a regular simplex. Here $\mathbf{A}_R = (\mathbf{A}')(\mathbf{AA}')^{-1} = \begin{bmatrix}1\\1\\1\end{bmatrix}\{[1\ 1\ 1]\begin{bmatrix}1\\1\\1\end{bmatrix}\}^{-1} = \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$ and $\tilde{\mathbf{X}}^0 = (\mathbf{A}')(\mathbf{AA}')^{-1}\mathbf{B} = \begin{bmatrix}4\\4\\4\end{bmatrix}$. The distance from $\tilde{\mathbf{X}}^0$ to any of the three vertices where the simplex meets an axis, at $\begin{bmatrix}12\\0\\0\end{bmatrix}, \begin{bmatrix}0\\12\\0\end{bmatrix}$ or $\begin{bmatrix}0\\0\\12\end{bmatrix}$, is 9.8. Similarly, the distance from $\tilde{\mathbf{X}}^0$ to the three sides of the simplex (on the $x_1 = 0$, $x_2 = 0$ or $x_3 = 0$ planes) is 4.9. Therefore, $\tilde{\mathbf{X}}^0$ is indeed at the center of the constraint.

9.7. For $\mathbf{M} = \begin{bmatrix}m_{11} & m_{12} & m_{13}\\ m_{21} & m_{22} & m_{23}\\ m_{31} & m_{32} & m_{33}\end{bmatrix}$, the general result is $(\ell_{33})^2 = m_{33} - (\ell_{31})^2 - (\ell_{32})^2$. It is straightforward algebra, best done on computer software, to find $\ell_{21} = m_{21}/(m_{11})^{0.5}$, $\ell_{22} = (|\mathbf{M}_2|/|\mathbf{M}_1|)$, $\ell_{31} = m_{31}/(m_{11})^{0.5}$, $\ell_{32} = (m_{32} - \ell_{21}\ell_{31})/\ell_{22}$, and, making all the substitutions, $(\ell_{33})^2 = |\mathbf{M}_3|/|\mathbf{M}_2|$.

Chapter 10

10.1. $\mathbf{X}^* = [0.5 \quad 1.5 \quad 0]'$.

10.2. $\mathbf{X}^* = \begin{bmatrix} 3/31 \\ 18/31 \end{bmatrix}$.

10.3. $\mathbf{X}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

10.4. (a) $L(\mathbf{X}, \mathbf{Y}) = -x_1^2 - x_2^2 - x_3^2 - y_1(x_1 + x_2 - 2) - y_2(2x_1 + 3x_2 - 12)$. From Problem 10.1, $\mathbf{X}^* = [0.5 \quad 1.5 \quad 0]'$; also from those results, using y's instead of λ's, we have $\mathbf{Y}^* = [3 \quad 0]'$.

(b) $L(\mathbf{X}, \mathbf{Y}) = -6x_1^2 - 5x_2^2 - y(-x_1 - 5x_2 + 3)$. $\mathbf{X}^* = \begin{bmatrix} 3/31 \\ 18/31 \end{bmatrix}$ and $y^* = 36/31$.

(c) $L(\mathbf{X}, \mathbf{Y}) = -6x_1^2 - 5x_2^2 - y(x_1 + 5x_2 - 3)$. $\mathbf{X}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $y^* = 0$.

10.5. (a) Problem 7-1. $\mathbf{X}^* = [0 \quad 0 \quad 3.6]'$, $\mathbf{S}^* = [8.4 \quad 0]'$, $\mathbf{Y}^* = [0 \quad 3]'$ and $\mathbf{T}^* = [-3 \quad -1 \quad 0]'$. For all of these problems with linear objective functions, $\nabla f(\mathbf{X}^*) = \nabla f(\mathbf{X})$. The constraints that are binding at the optimal solution are $h^2(\mathbf{X})$ along with the lower limits on x_1 and x_2, so $\mathbf{M}^* = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 5 & 0 & 0 \end{bmatrix}$, and $\nabla f(\mathbf{X}) = \begin{bmatrix} 3 \\ 2 \\ 15 \end{bmatrix} = (y_2^*)\nabla h^2(\mathbf{X}^*) + (t_1^*)\mathbf{I}_1 + (t_2^*)\mathbf{I}_2 = (3)\begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix} + (-3)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

(b) Problem 7-3. Here $\mathbf{X}^* = \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}$, $\mathbf{S}^* = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\mathbf{Y}^* = \begin{bmatrix} 0 \\ 2.25 \end{bmatrix}$ and $\mathbf{T}^* = \begin{bmatrix} 2.75 \\ 0 \end{bmatrix}$. $\nabla f(\mathbf{X}) = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$ and $\mathbf{M}^* = \begin{bmatrix} 1 & 1 \\ 4 & 0 \end{bmatrix}$ and indeed $\nabla f(\mathbf{X}) = (y_2^*)\nabla h^2(\mathbf{X}^*) + (t_1^*)\mathbf{I}_1 = (2.25)\begin{bmatrix} 1 \\ 4 \end{bmatrix} + (2.75)\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(c) Problem 7-4. $\mathbf{X}^* = [0 \quad 5 \quad 10 \quad 0]'$, $\mathbf{S}^* = [0 \quad 0]'$, $\mathbf{Y}^* = [2 \quad 4]'$ and $\mathbf{T}^* = [-3 \quad 0 \quad 0 \quad -11]'$. $\nabla f(\mathbf{X}) = [1 \quad 8 \quad 10 \quad 9]'$ and $\mathbf{M}^* = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 5 & 0 & 1 \end{bmatrix}$, so $\nabla f(\mathbf{X}) = (y_1^*)\nabla h^1(\mathbf{X}^*) + (y_2^*)\nabla h^2(\mathbf{X}^*) + (t_1^*)\mathbf{I}_1 + (t_4^*)\mathbf{I}_4 = (2)\mathbf{M}_1^* + (4)\mathbf{M}_2^* + (-3)\mathbf{M}_3^* + (-11)\mathbf{M}_4^*$.

10.6. (a) Example 1. $\mathbf{M}^* = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $\rho(\mathbf{M}^*) = 1$. (b) Example 2. $\mathbf{M}^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\rho(\mathbf{M}^*) = 1$. (c) Example 3. $\mathbf{M}^* = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}$ and

$\rho(M^*) = 2$. (d) Example 4. $M^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\rho(M^*) = 0 \neq 1$. (e) Example 5. $M^* = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\rho(M^*) = 1 \neq 2$. (f) Example 6. $M^* = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$ and $\rho(M^*) = 1 \neq 2$.

10.7. (a) $X^* = \begin{bmatrix} 20 \\ 0 \end{bmatrix}$. (b) $X^* = \begin{bmatrix} 12 \\ 0 \end{bmatrix}$. (c) $X^* = \begin{bmatrix} 24 \\ 6 \end{bmatrix}$.

Chapter 11

11.1. (a) The Hessian matrix for $f(\mathbf{X})$ in the primal problem is $\mathbf{H} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$, which is negative definite, so $f(\mathbf{X})$ is concave, and with linear constraints, the conditions for (11-11) to be the dual to (11-7) are met. Here this means that the dual nonlinear programming problem is

Minimize $\Delta(\mathbf{X}, \mathbf{Y}) = -x_1^2 - x_2^2 - x_3^2 + 4x_1 + 6x_2 - y_1(x_1 + x_2 - 2) - y_2(2x_1 + 3x_2 - 12) - x_1(-2x_1 + 4 - y_1 - 2y_2) - x_2(-2x_2 + 6 - y_1 - 3y_2) - x_3(-2x_3)$, subject to $y_1 + 2y_2 + 2x_1 - 4 \ge 0$, $y_1 + 3y_2 + 2x_2 - 6 \ge 0$, $2x_3 \ge 0$ and $y_1 \ge 0$, $y_2 \ge 0$.

(b) Here and in part (c), you can rewrite the nonlinear primal problem as one of maximization of $-f(\mathbf{X})$ and, in part (b), reverse the direction of the primal inequality by multiplying through by -1. Then the dual will be a minimization problem, formulated as in the text. In that case, $\mathbf{H} = \begin{bmatrix} -12 & 0 \\ 0 & -10 \end{bmatrix}$, so again we have a concave objective for the nonlinear primal maximization problem, with a linear constraint, and the structure of (11-11) indicates the dual. Alternatively, you could work out the nonlinear *maximization* problem that is dual to a nonlinear primal *minimization* problem. Using the former approach,

Minimize $\Delta(\mathbf{X}, y) = -6x_1^2 - 5x_2^2 - y(-x_1 - 5x_2 + 3) - x_1(-12x_1 + y) - x_2(-10x_2 + 5y)$, subject to $-y + 12x_1 \ge 0$, $-5y + 10x_2 \ge 0$ and $y \ge 0$.

[Here and in part (c), below, there is only one y since there is only one constraint in the primal problem.]

(c) Minimize $\Delta(\mathbf{X}, y) = -6x_1^2 - 5x_2^2 - y(x_1 + 5x_2 - 3) - x_1(-12x_1 - y) - x_2(-10x_2 - 5y)$, subject to $y + 12x_1 \ge 0$, $5y + 10x_2 \ge 0$ and $y \ge 0$.

11.2. (a) Starting with Problem 11.1(b), because it is smaller; after some algebra, the objective function is $\Delta(\mathbf{X}, y) = 6x_1^2 + 5x_2^2 - 3y$. Using μ's for the Lagrange multipliers in this minimization problem, the optimal solution is $\mu_1^* = 3/31 = x_1^*$, $\mu_2^* = 18/31 = x_2^*$ and $y^* = 36/31 = \lambda^*$. Slacks in the two dual constraints are $t_1^* = 0$ and $t_2^* = 0$; in the primal problem, $s^* = 0$.

The Hessian matrix for this problem is, $\mathbf{H}(\Delta) = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

This is only positive *semi*-definite, from the rules on all seven principal minors, given in (1-41) in Chapter 1. $\Delta(\mathbf{X}, y)$ does satisfy the *necessary* conditions for quasi-convexity, in (3-1-11) in Appendix 3.1.

The conditions in Theorem P-D3′ are $\mathbf{X}^* \circ \mathbf{T}^* = 0$ and $\mathbf{Y}^* \circ \mathbf{S}^* = 0$; it is clear that they are met by this pair of optimal solutions. Concerning Theorem P-D4′, note that $f(\mathbf{X}^*) = 1.742$ in the primal, where $\mathbf{X}^* = \begin{bmatrix} 3/31 \\ 18/31 \end{bmatrix} = \begin{bmatrix} 0.097 \\ 0.581 \end{bmatrix}$. Increasing the right-hand side of the primal constraint by 0.1 leads to $\mathbf{X}^{**} = \begin{bmatrix} 0.1 \\ 0.6 \end{bmatrix}$ and $f(\mathbf{X}^{**}) = 1.860$. Note that $y^* = 1.16$, and $f(\mathbf{X}^{**}) \cong f(\mathbf{X}^*) + (y^*)(\Delta b) = 1.742 + 0.116 = 1.858 \cong 1.86$.

(b) Here the simplified objective function is $\Delta(\mathbf{X}, \mathbf{Y}) = x_1^2 + x_2^2 + x_3^2 + 2y_1 + 12y_2$. The full solution for the nonlinear dual is $\mathbf{M}^* = [\mu_1^* \;\; \mu_2^* \;\; \mu_3^*]' = [0.5 \;\; 1.5 \;\; 0]'$, $\mathbf{Y}^* = [3 \;\; 0]'$, $\mathbf{T}^* = [t_1^* \;\; t_2^* \;\; t_3^*] = \mathbf{0}$. The optimal solution to the nonlinear primal was $\mathbf{X}^* = [0.5 \;\; 1.5 \;\; 0]'$, $\mathbf{S}^* = [0 \;\; 6.5]'$ and $\Lambda^* = [3 \;\; 0]'$. As in part (a), $\Delta(\mathbf{X}, \mathbf{Y})$ is "almost" quasiconvex, with a multidimensional trough-like shape, and the solution represents a minimum.

It is apparent that the conditions of Theorem P-D3′ are met. Also, in the primal problem, $f(\mathbf{X}^*) = 8.5$. If we increase b_1 by 0.1, so the constraint reads $x_1 + x_2 = 2.1$, $\mathbf{X}^{**} = \begin{bmatrix} 0.55 \\ 1.55 \end{bmatrix}$ and $f(\mathbf{X}^{**}) = 8.795 \cong f(\mathbf{X}^*) + y_1^*(\Delta b_1) = 8.5 + (3)(0.1) = 8.8$. Since the second constraint is not binding, increasing b_2 does nothing to the optimal solution; $y_2^* = 0$. The conditions of Theorem P-D4′ are satisfied.

11.3. (a) Problem 10.1 is

$$\begin{aligned} \text{Maximize} \quad & f(\mathbf{X}) = -x_1^2 - x_2^2 - x_3^2 + 4x_1 + 6x_2 \\ \text{subject to} \quad & x_1 + x_2 \le 2 \\ & 2x_1 + 3x_2 \le 12 \\ & \text{and } \mathbf{X} \ge \mathbf{0} \end{aligned}$$

(a1) Zoutendijk. Let $\mathbf{X}^0 = [0 \quad 0 \quad 0]'$. Since $\nabla f(\mathbf{X}^0) = [4 \quad 6 \quad 0]'$, move through the feasible region, from the origin, until a constraint is hit. (The gradient itself is the first feasible direction.) It is obvious from the problem that $x_3^* = 0$, so in what follows we will deal with finding the optimal x_1 and x_2 only. In this and other applications of Zoutendijk's approach, if there are only two unknowns, it is very helpful if you draw the solution space geometry for each of the Zoutendijk linear programs. Using the normalizations $-1 \le d_1^k \le 1$ and $-1 \le d_2^k \le 1$ means that the feasible region is contained in a square with vertical sides at $d_1^k = \pm 1$ and horizontal sides at $d_2^k = \pm 1$.

Since $x_1 + x_2 \le 2$ is hit first, $\mathbf{X}^1 = \begin{bmatrix} 0.8 \\ 1.2 \end{bmatrix}$. Now $\nabla^1 = \begin{bmatrix} 2.4 \\ 3.6 \end{bmatrix}$ and so the next Zoutendijk linear program is

$$\begin{aligned} \text{Maximize} \quad & 2.4d_1^1 + 3.6d_2^1 \\ \text{subject to} \quad & -1 \le d_1^1 \le 1 \\ & -1 \le d_2^1 \le 1 \\ & d_1^1 + d_2^1 \le 0 \end{aligned}$$

(The last constraint reflects the fact that at $\mathbf{X}^1$ we have bumped into constraint 1.) It is clear from this constraint that at least one of the d's will have to be negative; so the simplex method with nonnegativity requirements needs to be modified.

Considering the geometry of this problem, the feasible region is the left lower triangle in the feasible square, and the optimal solution is $d_1^1 = -1$, $d_2^1 = 1$, so $\mathbf{X}^2 = \mathbf{X}^1 + \alpha_2\mathbf{D}^1 = \begin{bmatrix} 0.8 \\ 1.2 \end{bmatrix} + \alpha_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Doing a line search along this direction, it is easily established that the maximum $f(\mathbf{X})$ is at $\alpha_2 = 0.3$, so that $\mathbf{X}^2 = \begin{bmatrix} 0.5 \\ 1.5 \end{bmatrix}$. Since this is not at an endpoint of constraint 1, there is no incentive to look further. (If you wanted to continue building Zoutendijk linear programs, you would find that at $\mathbf{X}^2$ the new linear program requires maximization of $3d_1^2 + 3d_2^2$, which is parallel to the constraint $d_1^2 + d_2^2 \le 0$, so there are multiple optima, including $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$--which means don't move at all. Other optima send you back and forth along the constraint, always telling you to undo the step that you have just taken.)

(a2) Rosen. At $\mathbf{X}^1$, $\mathbf{A}^1 = [1 \;\; 1]$, so, from (11-2), $\mathbf{M}^1 = \{\mathbf{I} - (\mathbf{A}^1)'[\mathbf{A}^1(\mathbf{A}^1)']^{-1}\mathbf{A}^1\} = \{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}0.5 & 0.5\\0.5 & 0.5\end{bmatrix}\} = \begin{bmatrix}0.5 & -0.5\\-0.5 & 0.5\end{bmatrix}$, and $\mathbf{G}^1 = \mathbf{M}^1\nabla^1 = \begin{bmatrix}0.5 & -0.5\\-0.5 & 0.5\end{bmatrix}\begin{bmatrix}2.4\\3.6\end{bmatrix} = \begin{bmatrix}-0.6\\0.6\end{bmatrix}$. Therefore, $\mathbf{X}^2 = \mathbf{X}^1 + \alpha_2\mathbf{G}^1$ and, using a line search along $\mathbf{G}^1$ confirms that the maximum occurs when $\alpha_2 = 0.5$, so that $\mathbf{X}^2 = \begin{bmatrix}0.5\\1.5\end{bmatrix}$. You can be convinced that this is indeed optimal by looking for $\mathbf{G}^2$, the projected gradient at this point. Since $\mathbf{M}^2 = \mathbf{M}^1$, and since $\nabla^2 = \begin{bmatrix}3\\3\end{bmatrix}$, $\mathbf{G}^2 = \mathbf{M}^2\nabla^2 = \begin{bmatrix}0.5 & -0.5\\-0.5 & 0.5\end{bmatrix}\begin{bmatrix}3\\3\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$.

Problem 10.2 is

$$\begin{aligned}\text{Minimize} \quad & 6x_1^2 + 5x_2^2\\ \text{subject to} \quad & x_1 + 5x_2 \geq 3\\ & \text{and } \mathbf{X} \geq 0\end{aligned}$$

(a3) Zoutendijk. Let $\mathbf{X}^0 = \begin{bmatrix}1\\1\end{bmatrix}$, a point inside the feasible region for this minimization problem. $\nabla^0 = \begin{bmatrix}12\\10\end{bmatrix}$; since this is a minimization problem, we move in the direction of the *negative* gradient. Starting out in this way, move along the negative gradient until the constraint is hit. This is at $\mathbf{X}^1 = \begin{bmatrix}3/26\\15/26\end{bmatrix}$, where $\nabla^1 = \begin{bmatrix}36/26\\150/26\end{bmatrix}$. Again, since direction is all that is important, normalize the gradient to $\nabla^1 = \begin{bmatrix}6\\25\end{bmatrix}$. The first Zoutendijk linear program is

$$\begin{aligned}\text{Maximize} \quad & -6d_1^1 - 25d_2^1\\ \text{subject to} \quad & -1 \leq d_1^1 \leq 1\\ & -1 \leq d_2^1 \leq 1\\ & d_1^1 + 5d_2^1 \leq 0\end{aligned}$$

From this (again it is easy if you sketch the feasible region), $\mathbf{D}^1 = \begin{bmatrix}-1\\0.2\end{bmatrix}$, so $\mathbf{X}^2 = \begin{bmatrix}3/26\\15/26\end{bmatrix} + \alpha_1\begin{bmatrix}-1\\0.2\end{bmatrix}$. A line search along $\begin{bmatrix}-1\\0.2\end{bmatrix}$ (now for a minimum) leads to $\mathbf{X}^2 = \begin{bmatrix}3/31\\18/31\end{bmatrix}$ when $\alpha_1 = 0.0186$.

(a4) Rosen. The route to $\mathbf{X}^1 = \begin{bmatrix} 3/26 \\ 15/26 \end{bmatrix}$ is the same as with Zoutendijk. Now $\mathbf{A}^1 = [1 \quad 5]$ so $\mathbf{M}^1 = \begin{bmatrix} 25/26 & -5/26 \\ -5/26 & 1/26 \end{bmatrix}$ and $\mathbf{G}^1 = \mathbf{M}^1\nabla^1 = \begin{bmatrix} 150/676 \\ -30/676 \end{bmatrix}$; the direction component is $\begin{bmatrix} 1 \\ -0.2 \end{bmatrix}$ and so the direction in $-\mathbf{G}^1 = \begin{bmatrix} -1 \\ 0.2 \end{bmatrix}$ is exactly the same as $\mathbf{D}^1$ in Zoutendijk's approach to this problem, and we will come to the same optimal point.

(b) The linear program in (7-6) is

$$\begin{aligned} \text{Maximize} \quad & \Pi = 2x_1 + 5x_2 \\ \text{subject to} \quad & x_1 + 2x_2 \le 10 \\ & 3x_1 + 2x_2 \le 24 \\ & x_1 + 10x_2 \le 40 \\ & \text{and } \mathbf{X} \ge \mathbf{0} \end{aligned}$$

(b1) Zoutendijk. Assume that $\mathbf{X}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$; for this problem $\nabla f(\mathbf{X}) = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, so move in this direction until a constraint is hit; in this problem, it is the third constraint 3. That point is $\mathbf{X}^1 = \begin{bmatrix} 20/13 \\ 50/13 \end{bmatrix}$. (It is not so important to identify each point in using either Rosen's or Zoutendijk's approach on a linear program because the gradient is constant, namely $\mathbf{P}'$.) Now determine a direction vector, $\mathbf{D}^1 = [d_1^1 \quad d_2^1]'$, from the linear program

$$\begin{aligned} \text{Maximize} \quad & 2d_1^1 + 5d_2^1 \\ \text{subject to} \quad & -1 \le d_1^1 \le 1 \\ & -1 \le d_2^1 \le 1 \\ & d_1^1 + 10d_2^1 \le 0 \end{aligned}$$

Again, as will generally be true in Zoutendijk linear programs, it is clear from the last constraint that at least one of the d's will have to be negative. From a simple sketch of the problem, $\mathbf{D}^1 = \begin{bmatrix} 1 \\ -0.1 \end{bmatrix}$.

Now we move along the original constraint, in the $\mathbf{D}^1$ direction, until the next constraint is met. This will be the first constraint, where $\mathbf{X} = \begin{bmatrix} 2.5 \\ 3.75 \end{bmatrix}$. So the next Zoutendijk linear program is

$$\begin{aligned} \text{Maximize} \quad & 2d_1^2 + 5d_2^2 \\ \text{subject to} \quad & -1 \le d_1^2 \le 1 \\ & -1 \le d_2^2 \le 1 \\ & d_1^2 + 10d_2^2 \le 0 \\ & d_1^2 + 2d_2^2 \le 0 \end{aligned}$$

It is easily established that $\mathbf{D}^2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, so there is no incentive to move from the current position, $\mathbf{X} = \begin{bmatrix} 2.5 \\ 3.75 \end{bmatrix}$, defined by the intersection of constraints 1 and 3 in the original problem.

(b2) Rosen. The initial step will be the same as with Zoutendijk, leading to $\mathbf{X}^1 = \begin{bmatrix} 20/13 \\ 50/13 \end{bmatrix}$ on constraint 3, so $\mathbf{A}^1 = [1 \quad 10]$ and, again using (11-2), $\mathbf{M}^1 = \begin{bmatrix} 100/101 & -10/101 \\ -10/101 & 1/101 \end{bmatrix}$ and so $\mathbf{G}^1 = \begin{bmatrix} 150/101 \\ -15/101 \end{bmatrix}$. What is important in $\mathbf{G}^1$ is *direction*, not magnitude, so "normalize" this to $\mathbf{G}^1 = \begin{bmatrix} 10 \\ -1 \end{bmatrix}$. Then $\mathbf{X}^2 = \mathbf{X}^1 + \alpha_2\mathbf{G}^1 = \begin{bmatrix} 20/13 \\ 50/13 \end{bmatrix} + \alpha_2\begin{bmatrix} 10 \\ -1 \end{bmatrix}$ and a line search indicates that we should go all the way to the next constraint (which is constraint 1), so $\mathbf{X}^2 = \begin{bmatrix} 2.5 \\ 3.75 \end{bmatrix}$. Since we are now at a corner and not on just an edge, as was $\mathbf{X}^1$, Rosen projects onto an edge. This is clear when you try to form $\mathbf{M}^2$ as in (11-2); $\mathbf{A}^2 = \begin{bmatrix} 1 & 2 \\ 1 & 10 \end{bmatrix}$ and $(\mathbf{A}^2)'[\mathbf{A}^2(\mathbf{A}^2)']^{-1}\mathbf{A}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ so $\mathbf{M}^2 = \mathbf{0}$. There are two possibilities:

(i) Let $\mathbf{A} = [1 \quad 10]$, so that we are investigating projecting onto constraint 3. (We know that we don't want to do this, since we just came from there, but it is useful to see how Rosen's method keeps us from doing it.) In that case, the "normalized" $\mathbf{G} = \begin{bmatrix} 10 \\ -1 \end{bmatrix}$, but moving in this direction is impossible; it presumes that we could still move rightward further along constraint 3, but that would take us outside of the feasible region, because we have bumped into constraint 1 also.

(ii) Let $\mathbf{A} = [1 \quad 2]$, so we are investigating a projection onto constraint 1. Now the normalized projected gradient is $G = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Movement in this direction is also impossible. It presumes that we can move leftward along constraint 1, which takes us outside of the feasible region.

The conclusion is don't move from $\mathbf{X}^2 = \begin{bmatrix} 2.5 \\ 3.75 \end{bmatrix}$.

11.4. (a) Zoutendijk. If you start at $\mathbf{X}^0 = 0$, where $\nabla^0 = [4 \quad 6 \quad 10]'$, you would not need to modify the initial gradient, and find $\mathbf{X}^1 = \mathbf{X}^0 + \alpha_1\nabla^0$, where you stopped when you hit the first constraint; this would be when $\alpha_1 = 0.4$ when $x_3 = 4$.

Alternatively, construct the initial Zoutendijk linear program at $\mathbf{X}^0$:

$$\begin{aligned} \text{Maximize} \quad & 4d_1^0 + 6d_2^0 + 10d_3^0 \\ \text{subject to} \quad & -1 \le d_1^0 \le 1 \\ & -1 \le d_2^0 \le 1 \\ & -1 \le d_3^0 \le 1 \\ & \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{D}^0 \le 0 \end{aligned}$$

The last three constraints reflect the fact that currently all x's are at their lower limits. From the structure of the problem, it is clear that $\mathbf{D}^0 = [1 \quad 1 \quad 1]'$, so $\mathbf{X}^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. A line search indicates that $f(\mathbf{X})$ is maximized when $\alpha_0 = 3.3$, so that $\mathbf{X}^1 = \begin{bmatrix} 3.3 \\ 3.3 \\ 3.3 \end{bmatrix}$. Here, $\nabla^1 = \begin{bmatrix} -2.6 \\ -0.6 \\ 3.4 \end{bmatrix}$, and the next linear program is

$$\begin{aligned} \text{Maximize} \quad & -2.6d_1^1 - 0.6d_2^1 + 3.4d_3^1 \\ \text{subject to} \quad & -1 \le d_1^1 \le 1 \\ & -1 \le d_2^1 \le 1 \\ & -1 \le d_3^1 \le 1 \end{aligned}$$

for which $\mathbf{D}^1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, so $\mathbf{X}^2 = \begin{bmatrix} 3.3 \\ 3.3 \\ 3.3 \end{bmatrix} + \alpha_1 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$. A line search for a maximum along this line takes us beyond $x_3 = 4$, so we stop when that constraint is hit ($\alpha_1 = 0.7$), giving $\mathbf{X}^2 = \begin{bmatrix} 2.6 \\ 2.6 \\ 4 \end{bmatrix}$.

Now $\nabla^2 = \begin{bmatrix} -1.2 \\ 0.8 \\ 2 \end{bmatrix}$ and the new Zoutendijk linear program, including the new constraint $d_3^2 \le 0$ (because we are at the upper limit of x_3), has an optimal solution $\mathbf{D}^2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, so $\mathbf{X}^3 = \begin{bmatrix} 2.6 \\ 2.6 \\ 4 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. The maximum for $f(\mathbf{X})$ along this direction is when $\alpha_2 =$ 0.5, giving $\mathbf{X}^3 = \begin{bmatrix} 2.1 \\ 3.1 \\ 4 \end{bmatrix}$. Here $\nabla^3 = \begin{bmatrix} -0.2 \\ -0.2 \\ 2 \end{bmatrix}$ and the solution of the next Zoutendijk linear program is $\mathbf{D}^3 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$ and, with an optimal $\alpha_3 = 0.1$, we find $\mathbf{X}^4 = \begin{bmatrix} 2.1 \\ 3.1 \\ 4 \end{bmatrix} + (0.1)\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$. At this point, $\nabla^4 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$, indicating that this is the optimal solution. You don't want to move away from the current x_1 or x_2, and you would like to increase x_3 but you can't, since it is already at its upper limit.

(b) Rosen. The route to $\mathbf{X}^1 = \begin{bmatrix} 3.3 \\ 3.3 \\ 3.3 \end{bmatrix}$ and then to $\mathbf{X}^2 = \begin{bmatrix} 2.6 \\ 2.6 \\ 4 \end{bmatrix}$ will be exactly the same as in Zoutendijk's approach. At $\mathbf{X}^2$, $\nabla^2 = \begin{bmatrix} -1.2 \\ 0.8 \\ 2 \end{bmatrix}$, and x_3 is at its upper limit, so $\mathbf{A}^2 = [0 \;\; 0 \;\; 1]$, $\mathbf{M}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{G}^2 = \begin{bmatrix} -1.2 \\ 0.8 \\ 0 \end{bmatrix}$. $\mathbf{X}^3 = \begin{bmatrix} 2.6 \\ 2.6 \\ 4 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1.2 \\ 0.8 \\ 0 \end{bmatrix}$, and a line search will turn up a maximum when $\alpha_2 = 0.5$, so that $\mathbf{X}^3 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$. At this point, $\nabla^3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$, and, as with Zoutendijk, we are finished.

11.5. (a) With $f_1 = 18 - 2(x_1 - 3)^2$, $f_2 = 49 - (x_2 - 7)^2$, $g_1(x_1) = x_1^2$ and $g_2(x_2) = x_2$ and given the mesh points indicated, we have

	x_1	$f_1(x_1)$	$g_1(x_1)$	x_2	$f_2(x_2)$	$g_2(x_2)$
x_{0j}	0	0	0	0	0	0
x_{1j}	1	10	1	2	13	1
x_{2j}	2	16	4	4	24	2

x_{3j}	3	18	9	6	33	3
x_{4j}	4	16	16	8	40	4

The separable programming problem is

$$\text{Maximize } 0\alpha_{01} + 10\alpha_{11} + 16\alpha_{21} + 18\alpha_{31} + 16\alpha_{41} + 0\alpha_{02} + 13\alpha_{12} + 24\alpha_{22} + 33\alpha_{32} + 40\alpha_{42}$$

subject to

$$0\alpha_{01} + \alpha_{11} + 4\alpha_{21} + 9\alpha_{31} + 16\alpha_{41} + 0\alpha_{02} + 1\alpha_{12} + 2\alpha_{22} + 3\alpha_{32} + 4\alpha_{42} \leq 16$$

$$\alpha_{01} + \alpha_{11} + \alpha_{21} + \alpha_{31} + \alpha_{41} \leq 1$$

$$\alpha_{02} + \alpha_{12} + \alpha_{22} + \alpha_{32} + \alpha_{42} \leq 1$$

and all α's nonnegative

In addition, you need the restricted basis entry rule; positive α_{ij} must be adjacent for j = 1 and for j = 2.

After several pivots, $\alpha^*_{31} = 1$ and $\alpha^*_{32} = 1$. In terms of the original problem, $x^*_1 = \sum_{k=0}^{5} \alpha^*_{k1}x_{k1} = (1)(3) = 3$ and $x^*_2 = \sum_{k=0}^{5} \alpha^*_{k2}x_{k2} = (1)(6) = 6$ and $f(\mathbf{X}^*) = 66$.

Note: the reason for using 18 in $f_1(x_1)$ and 49 in $f_2(x_2)$ was to create coefficients of zero on α_{01} and α_{02} in both the objective function and the first constraint. This, in turn, allowed us to bypass the two-phase approach for linear programs with equality constraints, since α_{01} plays the role of slack variable in the second constraint and α_{02} does the same in the third constraint.

(b) Ignoring nonnegativity and the constraint, $\mathbf{X}^* = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$. This satisfies both the constraint and nonnegativity and $f(\mathbf{X}^*) = 67$. Since the unconstrained maximum satisfies the constraint, it is unnecessary to also investigate the solution that occurs when the constraint is imposed as an equality.

11.6. For the original problem, only one pivot is required to reach the optimal solution, which is $\mathbf{X}^* = \begin{bmatrix} 0 \\ 10 \end{bmatrix}$ and $\Pi^* = 40$. For the rewritten version of the problem, after two pivots, $\mathbf{X}^* = \begin{bmatrix} 0 \\ 10 \\ 40 \end{bmatrix}$ and $\Pi^* = 40$. Note that $x^*_3 = \Pi^*$, as we would expect.

11.7. For the original problem, $\mathbf{X}^* = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ and $f(\mathbf{X}^*) = 120$.

For the rewritten problem, $\mathbf{X}^* = [x_1^* \;\; x_2^* \;\; x_3^*]' = [4 \;\; 6 \;\; 120]'$.

11.8. The transformed problem is

$$\text{Maximize} \quad F(x_1, \ldots, x_n, x_{n+1}) = 0x_1 + \ldots + 0x_n + x_{n+1}$$

$$\text{subject to } h^i(\mathbf{X}) \le 0$$

$$-f(\mathbf{X}) + x_{n+1} \le 0$$

$$\text{and } x_1 \ge 0, \ldots, x_n \ge 0$$

The Kuhn-Tucker conditions are:

$$(a) \begin{cases} F_j - \sum_{i=1}^{m} \lambda_i h_j^i + \lambda_{m+1} f_j \le 0 & (j = 1, \ldots, n) \\ 1 - \lambda_{m+1} \le 0 \end{cases}$$

$$(b) \begin{cases} x_j(F_j - \sum_{i=1}^{m} \lambda_i h_j^i + \lambda_{m+1} f_j) = 0 \\ x_{n+1}(1 - \lambda_{m+1}) = 0 \end{cases}$$

$$(c) \begin{cases} h^i(\mathbf{X}) \le 0 \\ -f(\mathbf{X}) + x_{n+1} \le 0 \end{cases}$$

$$(d) \begin{cases} \lambda_i[h^i(\mathbf{X})] = 0 \\ \lambda_{m+1}(-f(\mathbf{X}) + x_{n+1}) = 0 \end{cases}$$

$$(e) \; x_1 \ge 0, \ldots, x_n \ge 0, \; \lambda_1 \ge 0, \ldots, \lambda_{m+1} \ge 0$$

Notice that $F_j = 0$ $(j = 1, \ldots, n)$, so (a) and (b) become

$$(a) \begin{cases} -\sum_{i=1}^{m} \lambda_i h_j^i + \lambda_{m+1} f_j \le 0 & (j = 1, \ldots, n) \\ 1 - \lambda_{m+1} \le 0 \end{cases}$$

$$(b) \begin{cases} x_j(-\sum_{i=1}^{m} \lambda_i h_j^i + \lambda_{m+1} f_j) = 0 \\ x_{n+1}(1 - \lambda_{m+1}) = 0 \end{cases}$$

The Kuhn-Tucker conditions applied directly to the untransformed problem are:

$$(a') \quad f_j - \sum_{i=1}^{m} \lambda_i h_j^i \le 0$$

$$(b') \quad x_j(f_j - \sum_{i=1}^{m} \lambda_i h_j^i) = 0$$

$$(c') \quad h^i(\mathbf{X}) \le 0$$

$$(d') \quad \lambda_i[h^i(\mathbf{X})] = 0$$

(e′) $x_1 \geq 0, \ldots, x_n \geq 0, \lambda_1 \geq 0, \ldots, \lambda_m \geq 0$

Starting with (a) in the conditions for the transformed problem, $\lambda_{m+1} \geq 1$. From (d), this means that $x_{n+1} = f(\mathbf{X})$, always.

(i) If $\lambda_{m+1} = 1$, the conditions in (a) - (e) are exactly the same as those in (a′) - (e′).

(ii) If $\lambda_{m+1} > 1$, then from (d), $x_{n+1} = 0 = f(\mathbf{X})$. If you work through the implications in (b) and (d), you will find that exactly the same possibilities occur as you would find in (a′) - (e′).

11.9. Solution of this problem follows exactly the same steps as were used for Problem 11.8. However, in this case the transformed objective is to minimize x_{n+1} and the added constraint is $x_{n+1} \geq f(\mathbf{X})$.

11.10. (a) Here $\nabla h(\mathbf{X}) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$, and with $\mathbf{X}^0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $h(\mathbf{X}^0) = -8$. So the first constraint, c_1, to be used in conjunction with the linear objective function to create an approximating linear program is $-8 + [4 \quad 4]\begin{bmatrix} x_1 - 2 \\ x_2 - 2 \end{bmatrix} \leq 0$, which is $x_1 + x_2 \leq 6$. In conjunction with $f(\mathbf{X}) = 6x_1 + 8x_2$ and $\mathbf{X} \geq \mathbf{0}$, the optimal solution to this problem is $\mathbf{X}^1 = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$.

Now $\nabla h(\mathbf{X}^1) = \begin{bmatrix} 0 \\ 12 \end{bmatrix}$, $h(\mathbf{X}^1) = 20$ and the next linear constraint, c_2, is $20 + [0 \quad 12]\begin{bmatrix} x_1-0 \\ x_2-6 \end{bmatrix} \leq 0$, or $x_2 \leq 4.333$. The maximum is at the intersection of c_1 and c_2, so $X^2 = \begin{bmatrix} 1.67 \\ 4.33 \end{bmatrix}$.

Now $\nabla h(\mathbf{X}^2) = \begin{bmatrix} 3.33 \\ 8.67 \end{bmatrix}$, $h(\mathbf{X}^2) = 5.56$. Since 5.56 measures the amount by which the original constraint is violated by $\mathbf{X}^2$, and since this is greater than $\varepsilon = 1.0$, we continue. The next linear constraint, c_3, is $30x_1 + 78x_2 \leq 338$. The maximum is now at the intersection of c_1 and c_3, where $\mathbf{X}^3 = \begin{bmatrix} 2.71 \\ 3.29 \end{bmatrix}$.

$\nabla h(\mathbf{X}^3) = \begin{bmatrix} 5.42 \\ 6.58 \end{bmatrix}$, $h(\mathbf{X}^3) = 2.16$ and the next constraint, c_4, is $5.42x_1 + 6.58x_2 \leq 34.17$. The maximum to this linear program with now four constraints is at the extreme point defined by the intersection of c_3 and c_4; $\mathbf{X}^4 = \begin{bmatrix} 1.96 \\ 3.58 \end{bmatrix}$.

$\nabla h(\mathbf{X}^4) = \begin{bmatrix} 3.92 \\ 7.16 \end{bmatrix}$, $h(\mathbf{X}^4) = 0.66$ and the next constraint, c_5, is $3.91x_1 + 7.16x_2 \le 32.65$, for which $\mathbf{X}^5 = \begin{bmatrix} 2.29 \\ 3.31 \end{bmatrix}$, where c_4 and c_5 intersect. Since $h(\mathbf{X}^5) = 0.18 < \varepsilon$, stop. Using a smaller ε would carry us closer to the true solution, found in part (b), below.

(b) (i) Ignoring nonnegativity and without the constraint, the problem is meaningless, since $f(\mathbf{X})$ is linear, and so $f(\mathbf{X}) \to \infty$ as $x_1 \to \infty$ and $x_2 \to \infty$. (ii) With the constraint as an equation, the Lagrangian function is $L = 6x_1 + 8x_2 - \lambda(x_1^2 + x_2^2 - 16)$ and $x_1^* = \pm 2.4$ and $x_2^* = \pm 3.2$. The positive solutions must be chosen, so $\mathbf{X}^* = \begin{bmatrix} 2.4 \\ 3.2 \end{bmatrix}$, $\lambda^* = 3/x_1^* = 4/x_2^* = 1.25$ and $f(\mathbf{X}^*) = 40$.

11.11. Here is the series of cutting lines to create a series of linear programming problems, with $f(\mathbf{X}) = x_1 + x_2$, along with other relevant information at each solution:

k	Constraint	$\mathbf{X}^k$	$h(\mathbf{X}^k)$	$\nabla h(\mathbf{X}^k)$
0		$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	-7	$\begin{bmatrix} 6 \\ 8 \end{bmatrix}$
1	c_1 $6x_1 + 8x_2 \le 29$	$\begin{bmatrix} 4.83 \\ 0 \end{bmatrix}$	52.08	$\begin{bmatrix} 29 \\ 0 \end{bmatrix}$
2	c_2 $x_1 \le 3.03$	$\begin{bmatrix} 3.03 \\ 1.35 \end{bmatrix}$	13.19	$\begin{bmatrix} 18.18 \\ 5.4 \end{bmatrix}$
3	c_3 $18.18x_1 + 5.40x_2 \le 48.64$	$\begin{bmatrix} 2.06 \\ 2.08 \end{bmatrix}$	3.38	$\begin{bmatrix} 12.36 \\ 8.32 \end{bmatrix}$
4	c_4 $12.36x_1 + 8.32x_2 \le 39.39$	$\begin{bmatrix} 1.51 \\ 2.49 \end{bmatrix}$	1.24	$\begin{bmatrix} 9.06 \\ 9.96 \end{bmatrix}$
5	c_5 $9.06x_1 + 9.96x_2 \le 37.24$	$\begin{bmatrix} 1.74 \\ 2.16 \end{bmatrix}$	0.41	$\begin{bmatrix} 10.44 \\ 8.64 \end{bmatrix}$
6	c_6 $10.44x_1 + 8.64x_2 \le 36.42$	$\begin{bmatrix} 1.59 \\ 2.29 \end{bmatrix}$	0.07	$\begin{bmatrix} 9.54 \\ 9.16 \end{bmatrix}$
7	c_7 $9.54x_1 + 9.16x_2 \le 36.07$	$\begin{bmatrix} 1.508 \\ 2.367 \end{bmatrix}$	0.028	$\begin{bmatrix} 9.06 \\ 9.48 \end{bmatrix}$
8	c_8 $9.06x_1 + 9.48x_2 \le 36.074$	$\begin{bmatrix} 1.5445 \\ 2.3292 \end{bmatrix}$	0.007	so STOP

The optimal solution to this problem using exact methods (not approximations) is $\mathbf{X}^* = \begin{bmatrix} 1.5492 \\ 2.3238 \end{bmatrix}$.

11.12. (a) Here are results for

$$L(\mathbf{X};\mu) = -6x_1 - 8x_2 - \mu \ln(-x_1^2 - x_2^2 + 16)$$

(ignoring explicit consideration of $\mathbf{X} \geq 0$).

μ	x_1	x_2	$f(\mathbf{X})$
1.0	2.3408	3.1210	39.0128
0.1	2.3940	3.1920	39.9000
0.01	2.3973	3.2008	39.9902

(b) The exact answer is $\mathbf{X}^* = \begin{bmatrix}2.4\\3.2\end{bmatrix}$.

11.13. Here are results for

$$L(\mathbf{X};\mu) = -x_1 - x_2 - \mu \ln(-3x_1^2 - 2x_2^2 + 18)$$

μ	x_1	x_2	$f(\mathbf{X})$
1.0	1.2	1.8	3.0
0.1	1.5097	2.2645	3.7742
0.01	1.5452	2.3178	3.8630
0.0001	1.5492	2.3237	3.8729

As we saw in the solution to Problem 11.11, the exact solution is $\mathbf{X}^* = \begin{bmatrix}1.5492\\2.3238\end{bmatrix}$, and $f(\mathbf{X}^*) = 3.873$.

11.14. Here are results for an exterior point SUMT approach:

$$L(\mathbf{X};\mu) = x_1 + x_2 - \mu \ln(3x_1^2 + 2x_2^2 - 18)$$

μ	x_1	x_2	$f(\mathbf{X})$
1.0	2.0	3.0	5.0
0.1	1.5897	2.3846	3.9743
0.01	1.5532	2.3298	3.8830
0.0001	1.5492	2.3239	3.8731

11.15. (a) For this quadratic program

$$\mathbf{M} = \begin{bmatrix}-4 & 0 & -1 & 1 & 0 & 0\\ 0 & -2 & -1 & 0 & 1 & 0\\ 1 & 1 & 0 & 0 & 0 & 1\end{bmatrix} \text{ and } \mathbf{N} = \begin{bmatrix}-12\\-14\\5\end{bmatrix}$$

and $\mathbf{V} = [x_1\ x_2\ \lambda_1\ \lambda_2\ \lambda_3\ s]'$. There are up to $C_3^6 = 20$ basic solutions to the system $\mathbf{MV} = \mathbf{N}$, as in (11-55). Of these, three have singular coefficient matrices and of the remaining 17, 12 generate solutions that fail nonnegativity requirements. Of the remaining five, four fail one of the conditions in (11-53")(b). The one that is left is $\mathbf{X}^* = \begin{bmatrix}1.333\\3.667\end{bmatrix}$, which is the correct solution.

(b) For the current problem, $g_1(x_1) = x_1$ and $g_2(x_2) = x_2$. Using mesh points 0, 1, 2, 3, 4, 5,

	x_j (j = 1, 2)	$f_1(x_1)$	$g_1(x_1)$	$f_2(x_2)$	$g_2(x_2)$
x_{0j}	0	0	0	0	0
x_{1j}	1	10	1	13	1
x_{2j}	2	16	2	24	2
x_{3j}	3	18	3	33	3
x_{4j}	4	16	4	40	4
x_{5j}	5	10	5	45	5

The separable programming problem is

$$\text{Maximize } 0\alpha_{01} + 10\alpha_{11} + 16\alpha_{21} + 18\alpha_{31} + 16\alpha_{41} + 10\alpha_{51} + 0\alpha_{02} + 13\alpha_{12} + 24\alpha_{22} + 33\alpha_{32} + 40\alpha_{42} + 45\alpha_{52}$$

subject to

$$0\alpha_{01} + \alpha_{11} + 2\alpha_{21} + 3\alpha_{31} + 4\alpha_{41} + 5\alpha_{51} + 0\alpha_{02} + 1\alpha_{12} + 2\alpha_{22} + 3\alpha_{32} + 4\alpha_{42} + 5\alpha_{52} \le 5$$

$$\alpha_{01} + \alpha_{11} + \alpha_{21} + \alpha_{31} + \alpha_{41} + \alpha_{51} \le 1$$

$$\alpha_{02} + \alpha_{12} + \alpha_{22} + \alpha_{32} + \alpha_{42} + \alpha_{52} \le 1$$

and all α's nonnegative

In addition, there is the restricted basis entry rule; positive α_{ij} must be adjacent for j = 1 and for j = 2.

After several pivots, $\alpha^*_{11} = 1$ and $\alpha^*_{42} = 1$. In terms of the original problem, this means $x^*_1 = \sum_{k=0}^{5} \alpha^*_{k1}x_{k1} = (1)(1) = 1$ and $x^*_2 = \sum_{k=0}^{5} \alpha^*_{k2}x_{k2} = (1)(4) = 4$ and $f(\mathbf{X}^*) = 50$.

(c) $\mathbf{X}^* = \begin{bmatrix} 1.333 \\ 3.667 \end{bmatrix}$ and $f(\mathbf{X}^*) = 50.333$.

11.16. For the primal maximization problem, $\mathbf{H} = \begin{bmatrix} -4 & 0 \\ 0 & -2 \end{bmatrix}$, so $f(\mathbf{X})$ is concave and the constraint is linear. Then the structure of the dual problem is as given in (11-11), namely

$$\text{Minimize } \Delta(\mathbf{X}, y) = -2x_1^2 - x_2^2 + 12x_1 + 14x_2 - y(x_1 + x_2 - 5) - x_1(-4x_1 + 12 - y) - x_2(-2x_2 + 14 - y)$$

subject to $y + 4x_1 \ge 12$

$y + 2x_2 \ge 14$

and $y \ge 0$

After algebra, this objective function is

$$\text{Minimize } \Delta(\mathbf{X}, y) = 2x_1^2 + x_2^2 + 5y$$

Using the full set of Kuhn-Tucker conditions on this problem requires examination of eight alternative possibilities. All but one will be rejected, leaving $\mathbf{X}^* = \begin{bmatrix} 4/3 \\ 11/3 \end{bmatrix}$, $y^* = 20/3$ and $f(\mathbf{X}^*) = 50.333$. As was the case with the minimization functions in Problem 11.2, $\Delta(\mathbf{X}, y)$ has a trough-like shape and fails sufficiency tests for convexity or quasiconvexity, but $\mathbf{X}^*$, y^* represents a constrained minimum point.

Theorem P-D3′ requires $y^*(x_1^* + x_2^* - 5) = 0$, $x_1^*(-4x_1^* + 12 - y^*) = 0$ and $x_2^*(-2x_2^* + 14 - y^*) = 0$. It is easily verified that all three of these conditions are satisfied.

Theorem P-D4′ indicates that $\partial\Pi(\mathbf{X}^*)/\partial b_i = y_i^*$ under certain conditions. If you redo the primal problem, with the constraint changed to $x_1 + x_2 \le 6$, $\mathbf{X}^* = \begin{bmatrix} 5/3 \\ 13/3 \end{bmatrix}$ and $f(\mathbf{X}^*) = 56.333$. The change in $f(\mathbf{X}^*)$, for a unit *increase* in the right-hand side of the constraint, is 6. Now consider the primal with the constraint changed to $x_1 + x_2 \le 4$; the solution is $\mathbf{X}^* = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $f(\mathbf{X}^*) = 43$. The change in $f(\mathbf{X}^*)$, for a unit *decrease* in the right-hand side of the constraint, is -7.333. The *average* change, for a unit increase or a unit decrease, is 6.667 and $y^* = 6.667$, as the theorem suggests.