Exercises

13.3.1 By multiplying out the brackets in the expression $(\lambda \mathbf{a} + \mu \mathbf{b})^{\mathrm{T}}(\lambda \mathbf{a} + \mu \mathbf{b})$, prove L3.

13.3.2 Show that, if $\mathbf{a}$ and $\mathbf{b}$ are n–vectors such that $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$, then $-1 \leq \mathbf{a}^{\mathrm{T}}\mathbf{b} \leq 1$. Give examples where $n = 2$, $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$ and

$$\text{(a) } \mathbf{a}^{\mathrm{T}}\mathbf{b} = 1; \quad \text{(b) } \mathbf{a}^{\mathrm{T}}\mathbf{b} = -1; \quad \text{(c) } \mathbf{a}^{\mathrm{T}}\mathbf{b} = 0; \quad \text{(d) } 0 < \mathbf{a}^{\mathrm{T}}\mathbf{b} < 1.$$

13.3.3 Prove the following n–dimensional version of Pythagoras' theorem: if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal n–vectors then

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

13.3.4 Prove that the determinant of an orthogonal matrix is either 1 or -1. Show that both cases can occur.

13.3.5 In Exercise 11.2.7, you were asked to find the 2×2 matrix that describes anti-clockwise rotation about the origin through 45 degrees. Verify that this is an orthogonal matrix.

13.3.6 Give an example of a 3×2 matrix $\mathbf{U}$ such that $\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_2$. Is it an orthogonal matrix?

13.3.7 Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}.$$

Verify that any two of the three 3–vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are orthogonal. Find scalars λ, μ, ν such that $[\lambda \mathbf{u} \;\; \mu \mathbf{v} \;\; \nu \mathbf{w}]$ is an orthogonal matrix.

13.4 Quadratic forms and symmetric matrices

In Chapter 4 we defined a quadratic function to be a function of the form

$$f(x) = ax^2 + bx + c$$

where a, b, c are constants. A slight generalisation of this is the expression

$$q(x, y) = ax^2 + bxy + cy^2,$$

which is known as a quadratic form in the two variables x and y. Notice that $q(x, y)$ reduces to $f(x)$ if we set $y = 1$.

By the rules of matrix multiplication, we may write

$$q(x, y) = \mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} \qquad (13.10)$$

where $\mathbf{z} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} a & s \\ t & c \end{bmatrix}$ and s and t are any two numbers such that $s + t = b$. The expression $q(\mathbf{z})$ has the same meaning as $q(x, y)$.

If we impose the condition that $s = t$, their common value must be $b/2$; then $\mathbf{A}$ is determined uniquely by the coefficients a, b, c of the quadratic form, as the matrix

$$\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}.$$

Imposing the condition that $s = t$ is the same as requiring that $\mathbf{A}$ be its own transpose; a matrix with this property is said to be symmetric. Generally, a **symmetric matrix** of order n is defined to be an $n \times n$ matrix $\mathbf{A}$ such that $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$. Notice that any diagonal matrix is symmetric; in particular, the identity matrix $\mathbf{I}$ is symmetric.

We have shown that, given a quadratic form in two variables $q(x, y)$, there is exactly one symmetric 2×2 matrix $\mathbf{A}$ satisfying (13.10). Now consider the general quadratic form in three variables:

$$q(x_1, x_2, x_3) = ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3,$$

where a, b, c, d, e, f are constants. Let $\mathbf{x} = [x_1 \;\; x_2 \;\; x_3]^{\mathrm{T}}$; then there is exactly one symmetric 3×3 matrix $\mathbf{A}$ such that $q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, namely

$$\begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix}.$$

All of this may be generalised. A **quadratic form** in n variables $x_1, x_2, \ldots, x_n$ is defined to be an expression of the form

$$q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x},$$

where $\mathbf{x} = [x_1 \;\; x_2 \;\; \ldots \;\; x_n]^{\mathrm{T}}$ and $\mathbf{A}$ is a symmetric $n \times n$ matrix. By requiring that $\mathbf{A}$ be symmetric, we ensure that $\mathbf{A}$ is uniquely determined by the coefficients of x_1^2, x_1x_2 and so on when the quadratic form is written out in full; we have just shown this for the special cases $n = 2$ and $n = 3$, and it is true for all n. The relation between quadratic forms and symmetric matrices is discussed in more detail in the appendix to this chapter.

Definite and semidefinite quadratic forms

When we discussed quadratic functions in Chapter 4, we devoted some attention to the question whether a function $f(x)$ was positive for all values of x. There is no direct analogue of this property for quadratic forms, since $q(\mathbf{0}) = 0$ for every quadratic form $q(\mathbf{x})$. However, given a quadratic form $q(\mathbf{x})$, we can ask whether $q(\mathbf{x}) > 0$ for every non-zero vector $\mathbf{x}$; if so, we say that $q(\mathbf{x})$ is a **positive definite** quadratic form.

Example 1 A quadratic form in n variables $x_1, \dots, x_n$ that is obviously positive definite is

$$x_1^2 + x_2^2 + \dots + x_n^2.$$

Recall that this is what we called $\|\mathbf{x}\|^2$ in the preceding section.

The quadratic form $q(\mathbf{x})$ is said to be **positive semidefinite** if $q(\mathbf{x}) \geq 0$ for every vector $\mathbf{x}$. Every positive definite quadratic form is positive semidefinite. On the other hand, it is easy to find quadratic forms which are positive semidefinite but not positive definite.

Example 2 Consider the quadratic form in three variables

$$q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 2x_2x_3.$$

This can be written as $x_1^2 + (x_2 - x_3)^2$ which, being a sum of squares, is non-negative. Therefore $q(x_1, x_2, x_3)$ is positive semidefinite; it is not positive definite because $q(0, 1, 1) = 0$.

We say that a quadratic form $q(\mathbf{x})$ is **negative semidefinite** if $-q(\mathbf{x})$ is positive semidefinite; this happens if and only if $q(\mathbf{x}) \leq 0$ for every vector $\mathbf{x}$. Similarly, $q(\mathbf{x})$ is said to be **negative definite** if $-q(\mathbf{x})$ is positive definite.

Many quadratic forms are neither positive semidefinite nor negative semidefinite.

Example 3 The quadratic form

$$q(x_1, x_2) = x_1^2 - x_2^2$$

is neither positive semidefinite nor negative semidefinite, since $q(0, 1) = -1$ and $q(1, 0) = 1$.

The terms positive definite, negative definite and so on are often applied to symmetric matrices as well as to the corresponding quadratic forms. Thus a symmetric matrix $\mathbf{A}$ is said to be positive definite if $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ whenever $\mathbf{x} \neq \mathbf{0}$, positive semidefinite if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x}$ and negative (semi)definite if $-\mathbf{A}$ is positive (semi)definite. For instance, the symmetric matrices associated with the quadratic forms of Examples 1, 2 and 3 are respectively

$$\mathbf{I}_n, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The first of these matrices is positive definite (for all n), the second is positive semidefinite but not positive definite and the third is neither positive semidefinite nor negative semidefinite.

Testing quadratic forms

Given a quadratic form, how do we test whether it is positive definite, negative definite or whatever? We can sometimes work from first principles as in the examples above. More typically, we use standard tests; these are most easily stated in terms of the associated symmetric matrix, but of course they translate immediately into results about the quadratic forms themselves. We begin with symmetric matrices of order 2.

(a) A 2×2 symmetric matrix is positive definite if and only if its diagonal entries are both positive and its determinant is positive.

(b) A 2×2 symmetric matrix is positive semidefinite if and only if its diagonal entries are both non-negative and its determinant is non-negative.

(c) A 2×2 symmetric matrix is negative definite if and only if its diagonal entries are both negative and its determinant is positive.

(d) A 2×2 symmetric matrix is negative semidefinite if and only if its diagonal entries are both non-positive and its determinant is non-negative.

Example 4 Let

$$\mathbf{A} = \begin{bmatrix} 2+t & 1 \\ 1 & 2-t \end{bmatrix}.$$

Then $\mathbf{A}$ is symmetric and $\det \mathbf{A} = 3 - t^2$. If $-\sqrt{3} < t < \sqrt{3}$, the determinant and both diagonal entries are positive: thus $\mathbf{A}$ is positive definite. If $t = \pm\sqrt{3}$, $\mathbf{A}$ is positive semidefinite but not positive definite. If $|t| > \sqrt{3}$, $\det \mathbf{A} < 0$, so $\mathbf{A}$ is neither positive semidefinite nor negative semidefinite.

Since $1 < \sqrt{3} < 2$, these results imply that the quadratic form

$$3x_1^2 + x_2^2 + 2x_1x_2$$

is positive definite, while the quadratic form $4x_1^2 + 2x_1x_2$ is neither positive semidefinite nor negative semidefinite.

We now discuss briefly why these tests for 2×2 symmetric matrices are valid. We focus on (a). Consider the symmetric matrix

$$\mathbf{A} = \begin{bmatrix} a & s \\ s & c \end{bmatrix}$$

and the associated quadratic form

$$q(x, y) = ax^2 + 2sxy + cy^2.$$

Then $a = q(1, 0)$ and $c = q(0, 1)$: for our matrix to be positive definite, we need these diagonal entries to be positive.

To see the relevance of the determinant, consider the case where $a = c = 1$. Then $\det \mathbf{A} = 1 - s^2$. To verify (a) in this case, we need to show that $q(x, y) > 0$ for all (x, y) other than $(0, 0)$ if and only if $s^2 < 1$. In fact,

$$q(x, y) = x^2 + 2sxy + y^2 = (x + sy)^2 + (1 - s^2)y^2$$

by completing the square. This is clearly positive if $s^2 < 1$ and at least one of x and y is not zero; while if $s^2 \geq 1$ we can make $q(x, y)$ non-positive by setting $x = s$, $y = -1$. Thus (a) holds for this matrix. The proofs of (a)–(d) in the general case consist of more elaborate versions of the same argument.

Higher dimensions

To generalise the tests (a)–(d) to matrices of order n, we need some definitions.

A **submatrix** of a matrix $\mathbf{A}$ is a matrix obtained from $\mathbf{A}$ by deleting some (or none) of its rows and some (or none) of its columns. A **principal submatrix** of a square matrix $\mathbf{A}$ is a submatrix obtained using the rule that the kth row of $\mathbf{A}$ is deleted if and only if the kth column of $\mathbf{A}$ is deleted. A **leading principal submatrix** of a square matrix $\mathbf{A}$ is a submatrix obtained by deleting the *last* m rows and columns, for some m.

For example, if

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ p & q & r \\ u & v & w \end{bmatrix},$$

then the principal submatrices of $\mathbf{A}$ are its three diagonal entries a, q, w (considered as 1×1 matrices), the three 2×2 matrices

$$\begin{bmatrix} a & b \\ p & q \end{bmatrix}, \quad \begin{bmatrix} a & c \\ u & w \end{bmatrix}, \quad \begin{bmatrix} q & r \\ v & w \end{bmatrix}$$

and $\mathbf{A}$ itself. The first, fourth and seventh of these matrices are the leading principal submatrices of $\mathbf{A}$.

A **minor** of a square matrix $\mathbf{A}$ is the determinant of a square submatrix of $\mathbf{A}$; a **principal minor** is the determinant of a principal submatrix and a **leading principal minor** is the determinant of a leading principal submatrix. In the 3×3 example just given, the leading principal minors of $\mathbf{A}$ are a, $aq - bp$ and $\det \mathbf{A}$; the other principal minors are q, w, $aw - cu$ and $qw - rv$.

Tests (a) and (b) for 2×2 symmetric matrices generalise as follows:

(a) An $n \times n$ symmetric matrix is positive definite if and only if its principal minors are all positive.

(b) An $n \times n$ symmetric matrix is positive semidefinite if and only if its principal minors are all non-negative.

These tests are usually cumbersome to apply. A more convenient version of (a) is the following:

(a′) An $n \times n$ symmetric matrix is positive definite if and only if its *leading* principal minors are all positive.

For example, the 3×3 symmetric matrix

$$\begin{bmatrix} f & 1 & 4 \\ 1 & g & 5 \\ 4 & 5 & h \end{bmatrix}$$

is positive definite if and only if the following three conditions are met: $f > 0$, $fg > 1$ and the determinant of the matrix is positive.

The test (a′) is often implemented in the following form:

(a″) An $n \times n$ symmetric matrix is positive definite if and only if it can be reduced by Gaussian elimination, *without row exchanges*, to an upper triangular matrix whose diagonal entries are all positive.

By contrast, non-negativity of the leading principal minors does not guarantee that a symmetric matrix is positive semidefinite; for example, the 3×3 diagonal matrix with diagonal entries $1, 0, -1$ has leading principal minors $1, 0, 0$ but is not positive semidefinite.

It is possible to generalise rules (c) and (d) to $n \times n$ symmetric matrices. However, it is usually easier to test for negative definiteness or semidefiniteness by applying (a′) or (b) to the matrix $-\mathbf{A}$.

Three propositions

We end this section with three general propositions about positive definite and semidefinite symmetric matrices. We give the proofs because they are good illustrations of the use of simple logical reasoning employing the algebra of inner products, but the reader may omit them without loss of continuity.

Proposition 1 Let $\mathbf{A}$ be a positive semidefinite symmetric matrix, $\mathbf{x}$ a vector. Then $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$ if and only if $\mathbf{A}\mathbf{x} \neq \mathbf{0}$.

Proof Let $\mathbf{y} = \mathbf{A}\mathbf{x}$. If $\mathbf{x}^{\mathrm{T}}\mathbf{y} > 0$, then obviously $\mathbf{y} \neq \mathbf{0}$. Now suppose $\mathbf{y} \neq \mathbf{0}$; we wish to show that $\mathbf{x}^{\mathrm{T}}\mathbf{y} > 0$. Since $\mathbf{A}$ is symmetric, $\mathbf{y}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}\mathbf{A}$; hence the expressions $\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ and $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y}$ are both equal to $\mathbf{y}^{\mathrm{T}}\mathbf{y}$. Therefore, for every scalar λ,

$$(\mathbf{x} - \lambda\mathbf{y})^{\mathrm{T}}\mathbf{A}(\mathbf{x} - \lambda\mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{y} - 2\lambda\mathbf{y}^{\mathrm{T}}\mathbf{y} + \lambda^2\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}.$$

Since $\mathbf{A}$ is positive semidefinite, the left-hand side of this equation is non-negative. Thus

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} \geq \lambda(2\mathbf{y}^{\mathrm{T}}\mathbf{y} - \lambda\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y})$$

for any real number λ. But since $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y}^{\mathrm{T}}\mathbf{y} > 0$. We may therefore choose λ to be positive but sufficiently small that $\lambda\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y} < 2\mathbf{y}^{\mathrm{T}}\mathbf{y}$. Hence $\mathbf{x}^{\mathrm{T}}\mathbf{y} > 0$.

Proposition 2 A symmetric matrix is positive definite if and only if it is positive semidefinite and invertible.

PROOF Let $\mathbf{A}$ be a positive semidefinite symmetric matrix. By Proposition 1, $\mathbf{A}$ is positive definite if and only if $\mathbf{Ax} \neq \mathbf{0}$ for every non-zero n–vector $\mathbf{x}$. The proposition now follows from the fact that a square matrix is invertible if and only if it is non-singular.

Proposition 3 If $\mathbf{A}$ is a positive definite symmetric matrix, so is $\mathbf{A}^{-1}$.

PROOF Let $\mathbf{A}$ be a positive definite symmetric matrix. By Proposition 2, $\mathbf{A}^{-1}$ exists and is invertible with inverse $\mathbf{A}$; it remains to show that $\mathbf{A}^{-1}$ is symmetric and positive semidefinite. Now $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$ for any invertible matrix $\mathbf{A}$; but in this case $(\mathbf{A}^{\mathrm{T}})^{-1} = \mathbf{A}^{-1}$ since $\mathbf{A}$ is symmetric; hence $(\mathbf{A}^{-1})^{\mathrm{T}} = \mathbf{A}^{-1}$, as required. To complete the proof, let $\mathbf{x}$ be any vector; we wish to show that $\mathbf{x}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{x} \geq 0$. Let $\mathbf{w} = \mathbf{A}^{-1}\mathbf{x}$. Since $\mathbf{A}^{-1}$ is symmetric, $\mathbf{w}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}\mathbf{A}^{-1}$. Therefore

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{x} = \mathbf{w}^{\mathrm{T}}\mathbf{x} = \mathbf{w}^{\mathrm{T}}\mathbf{A}\mathbf{w},$$

which is non-negative because $\mathbf{A}$ is positive definite.

Exercises

13.4.1 Show that, if $\mathbf{a}$ is a 2–vector, then $\mathbf{a}\mathbf{a}^{\mathrm{T}}$ is a 2×2 symmetric matrix. How do you think this result generalises?

13.4.2 A data set consists of n observations on two variables x_1 and x_2; the ith observation is denoted (x_{1i}, x_{2i}). Let $\mathbf{X}$ be the $n \times 2$ matrix whose ith row is $[x_{1i}\ \ x_{2i}]$.[4] Calculate the matrix $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, expressing its entries in $\sum$–notation, and verify that it is symmetric.

13.4.3 Prove that if $\mathbf{A}$ is a symmetric matrix, so is $\mathbf{B}^{\mathrm{T}}\mathbf{AB}$.

13.4.4 Show directly from the definitions that the quadratic form

$$q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_2 x_3$$

is positive definite. Write down the associated symmetric matrix.

13.4.5 Determine the symmetric matrix $\mathbf{A}$ such that, for every 2–vector $\mathbf{x}$,

$$2x_1^2 + 3x_2^2 + 4x_1x_2 = \mathbf{x}^{\mathrm{T}}\mathbf{Ax}$$

Use the test given in the text to show that the matrix $\mathbf{A}$ is positive definite.

13.4.6 Determine the values of t for which the symmetric matrix $\begin{bmatrix} 2t & 2 \\ 2 & t \end{bmatrix}$ is

[4] This notation, which is standard in the statistical literature, departs from the usual one for matrices, in that the second subscript indicates the row and the first the column.

(a) positive definite;
(b) positive semidefinite but not positive definite;
(c) negative definite;
(d) negative semidefinite but not negative definite;
(e) none of the above.

13.4.7 Determine the definiteness of the symmetric matrices

$$\begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 2 \\ 1 & 2 & 9 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 2 \\ 1 & 2 & 6 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

13.4.8 Suppose $\mathbf{C}$ is an $n \times k$ matrix. Show that $\mathbf{C}^{\mathrm{T}}\mathbf{C}$ is a positive semidefinite symmetric matrix. Show also that $\mathbf{C}^{\mathrm{T}}\mathbf{C}$ is positive definite if the columns of $\mathbf{C}$ are linearly independent. What restriction on n and k does this condition impose?

Problems on Chapter 13

13–1. Let G be the parallelogram in the xy–plane whose vertices have coordinates $(0,0)$, (a,b), (c,d) and $(a+c,\ b+d)$. Suppose that a, b, c, d are all positive. Show that the area of G is $|ad - bc|$ by enclosing G in a rectangle and subtracting the surrounding area.

13–2. (i) Let

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}.$$

Suppose $\mathbf{B}$ is non-singular and $\mathbf{A}$ is singular. In Problem 12–3 you were asked to prove that any change in the $(3,3)$ entry transforms $\mathbf{A}$ into an invertible matrix. Give another proof of the same result, by expanding $\det \mathbf{A}$ by its third row.

(ii) Sketch a proof of the following proposition: any singular square matrix can be transformed into an invertible matrix by arbitrarily small changes to its diagonal entries.
[HINT First prove the proposition for 2×2 matrices. Then use the argument of (i) to extend the proposition to 3×3 matrices. Finally, explain how the same logic may be used to prove the proposition for square matrices of higher order.]

(iii) Explain why the proposition of (ii) ceases to be true when the words 'singular' and 'invertible' are interchanged.

[This problem shows that invertible matrices can be considered as 'normal' square matrices and singular matrices as 'peculiar' ones. For this reason, it is common in applications to assume that square matrices are invertible unless one has a good reason to suppose otherwise.]

13–3. This problem develops further the n–good input-output model of Problems 11–4 and 12–4. Notation is as in those problems.

Suppose each industry j uses as inputs not only the produced goods $1, \ldots, n$ but also non-produced goods such as labour and raw materials; let the cost of such inputs, per unit of gross output of j, be c_j. Let p_j be the price of good j. Write down an expression for the cost of producing each unit of gross output of good j. Derive a system of linear equations which must hold if all industries exactly break even (price equals average cost).

Let $\mathbf{c}$ be the vector with components $c_1, \ldots, c_n$ and let $\mathbf{p}$ be the **price vector**, with components $p_1, \ldots, p_n$. What properties must the input-output matrix $\mathbf{A}$ have if, for every vector $\mathbf{c}$ with non-negative components, there is a corresponding vector $\mathbf{p}$ of non-negative prices such that all industries break even? How are these properties related to the answer to Problem 12–4?

13–4. This problem is concerned with the most popular of all econometric techniques, **least-squares estimation.**

Suppose you have a data set consisting of n observations on three variables y, x_1, x_2; the ith observation is denoted $[y_i \ \ x_{1i} \ \ x_{2i}]$. You wish to find a linear function of the form

$$y = b_1 x_1 + b_2 x_2$$

which fits the data as well as possible in the following sense: b_1 and b_2 are chosen so as to minimise the expression

$$Q(b_1, b_2) = \sum_{i=1}^{n} (y_i - b_1 x_{1i} - b_2 x_{2i})^2.$$

Let $\mathbf{y}$ be the n–vector whose ith component is y_i, $\mathbf{X}$ the $n \times 2$ matrix whose ith row is $[x_{1i} \ \ x_{2i}]$.[5] Assume that the columns of $\mathbf{X}$ are linearly independent. Let $\mathbf{b}$ be the 2–vector $[b_1 \ \ b_2]^{\mathrm{T}}$.

(i) Show that $Q(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\mathbf{b})$.

(ii) Suppose that $\mathbf{b}^*$ is a 2–vector such that

$$\mathbf{X}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\mathbf{b}^*) = \mathbf{0}. \qquad (*)$$

Using the result of Exercise 13.3.2, show that

$$Q(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b}^*)^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\mathbf{b}^*) + (\mathbf{b}^* - \mathbf{b})^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}(\mathbf{b}^* - \mathbf{b}).$$

(iii) Since the columns of $\mathbf{X}$ are linearly independent, the symmetric 2×2 matrix $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is positive definite (Exercise 13.4.8) and therefore invertible (Proposition 2 of Section 13.4). Deduce that there is only one vector $\mathbf{b}^*$ satisfying $(*)$, and find an explicit expression for $\mathbf{b}^*$.

[5] We commented on this notation in an earlier footnote, in connection with Exercise 13.4.2.

(iv) Show that $Q(\mathbf{b})$ is minimised when $\mathbf{b} = \mathbf{b}^*$.

[All of this generalises easily to the case where the data set consists of n observations on $1 + k$ variables $y, x_1, \ldots, x_k$. In particular, the case of three variables y, x_1, x_2, where the function to be fitted is of the form

$$y = b_1 x_1 + b_2 x_2 + b_3,$$

may be treated in this framework by letting $k = 3$, $x_{3i} = 1$ for all i.]