Numerical Analysis Lecture Notes

Peter J. Olver

12. Minimization

In this part, we will introduce and solve the most basic mathematical optimization problem: minimize a quadratic function depending on several variables. This will require a short introduction to positive definite matrices. Assuming the coefficient matrix of the quadratic terms is positive definite, the minimizer can be found by solving an associated linear algebraic system. With the solution in hand, we are able to treat a wide range of applications, including least squares fitting of data, interpolation, as well as the finite element method for solvilng boundary value problems for differential equations.

12.1. Positive Definite Matrices.

Minimization of functions of several variables relies on an extremely important class of symmetric matrices.

Definition 12.1. An $n \times n$ matrix K is called *positive definite* if it is symmetric, $K^T = K$, and satisfies the positivity condition

$$\mathbf{x}^T K\,\mathbf{x} > 0 \qquad \text{for all vectors} \qquad \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n. \tag{12.1}$$

We will sometimes write $K > 0$ to mean that K is a symmetric, positive definite matrix.

Warning: The condition $K > 0$ does *not* mean that all the entries of K are positive. There are many positive definite matrices that have some negative entries; see Example 12.2 below. Conversely, many symmetric matrices with all positive entries are not positive definite!

Remark: Although some authors allow non-symmetric matrices to be designated as positive definite, we will *only* say that a matrix is positive definite when it is symmetric. But, to underscore our convention and remind the casual reader, we will often include the superfluous adjective “symmetric” when speaking of positive definite matrices.

Given any symmetric matrix K, the homogeneous quadratic polynomial

$$q(\mathbf{x}) = \mathbf{x}^T K\,\mathbf{x} = \sum_{i,j=1}^{n} k_{ij}\, x_i\, x_j, \tag{12.2}$$

is known as a *quadratic form* on $\mathbb{R}^n$. The quadratic form is called *positive definite* if

$$q(\mathbf{x}) > 0 \qquad \text{for all} \qquad 0 \neq \mathbf{x} \in \mathbb{R}^n. \tag{12.3}$$

Thus, a quadratic form is positive definite if and only if its coefficient matrix is.

Example 12.2. Even though the symmetric matrix $K = \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix}$ has two negative entries, it is, nevertheless, a positive definite matrix. Indeed, the corresponding quadratic form

$$q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x} = 4x_1^2 - 4x_1x_2 + 3x_2^2 = (2x_1 - x_2)^2 + 2x_2^2 \geq 0$$

is a sum of two non-negative quantities. Moreover, $q(\mathbf{x}) = 0$ if and only if both $2x_1 - x_2 = 0$ and $x_2 = 0$, which implies $x_1 = 0$ also. This proves $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, and hence K is indeed a positive definite matrix.

On the other hand, despite the fact that $K = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ has all positive entries, it is *not* a positive definite matrix. Indeed, writing out

$$q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x} = x_1^2 + 4x_1x_2 + x_2^2,$$

we find, for instance, that $q(1, -1) = -2 < 0$, violating positivity. These two simple examples should be enough to convince the reader that the problem of determining whether a given symmetric matrix is or is not positive definite is not completely elementary.

Example 12.3. By definition, a general symmetric 2×2 matrix $K = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is positive definite if and only if the associated quadratic form satisfies

$$q(\mathbf{x}) = a x_1^2 + 2b x_1 x_2 + c x_2^2 > 0 \qquad \text{for all} \qquad \mathbf{x} \neq \mathbf{0}. \tag{12.4}$$

Analytic geometry tells us that this is the case if and only if

$$a > 0, \qquad\qquad a c - b^2 > 0, \tag{12.5}$$

i.e., the quadratic form has positive leading coefficient and positive determinant (or negative discriminant).

A practical test of positive definiteness comes from the following result, whose proof is based on Gaussian Elimination, [**42**].

Theorem 12.4. *A symmetric matrix K is positive definite if and only if it is regular and has all positive pivots.*

In other words, a square matrix K is positive definite if and only if it can be factored $K = L D L^T$, where L is special lower triangular and D is diagonal with all positive diagonal entries. Indeed, we cna then write the associated quadratic form as a sum of squares

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^T K \mathbf{x} = \mathbf{x}^T L D L^T \mathbf{x} = (L^T \mathbf{x})^T D \, (L^T \mathbf{x}) \\ &= \mathbf{y}^T D \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2, \end{aligned} \qquad \text{where} \qquad \mathbf{y} = L^T \mathbf{x}. \tag{12.6}$$

Furthermore, the resulting diagonal quadratic form is positive definite, $\mathbf{y}^T D \mathbf{y} > 0$ for all $\mathbf{y} \neq \mathbf{0}$ if and only if all the pivots are positive, $d_i > 0$. Invertibility of L^T tells us that $\mathbf{y} = \mathbf{0}$ if and only if $\mathbf{x} = \mathbf{0}$, and hence, positivity of the pivots is equivalent to positive definiteness of the original quadratic form: $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$.

Example 12.5. Consider the symmetric matrix $K = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 9 \end{pmatrix}$. Gaussian Elimination produces the factors

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}, \qquad L^T = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

in its factorization $K = L D L^T$. Since the pivots — the diagonal entries $1, 2$ and 6 in D — are all positive, Theorem 12.4 implies that K is positive definite, which means that the associated quadratic form satisfies

$$q(\mathbf{x}) = x_1^2 + 4 x_1 x_2 - 2 x_1 x_3 + 6 x_2^2 + 9 x_3^2 > 0, \qquad \text{for all} \qquad \mathbf{x} = (\, x_1, x_2, x_3 \,)^T \neq \mathbf{0}.$$

Indeed, the $L D L^T$ factorization implies that $q(\mathbf{x})$ can be explicitly written as a sum of squares:

$$q(\mathbf{x}) = x_1^2 + 4 x_1 x_2 - 2 x_1 x_3 + 6 x_2^2 + 9 x_3^2 = y_1^2 + 2 y_2^2 + 6 y_3^2, \tag{12.7}$$

where

$$y_1 = x_1 + 2 x_2 - x_3, \qquad y_2 = x_2 + x_3, \qquad y_3 = x_3,$$

are the entries of $\mathbf{y} = L^T \mathbf{x}$. Positivity of the coefficients of the y_i^2 (which are the pivots) implies that $q(\mathbf{x})$ is positive definite.

Slightly more generally, a quadratic form and its associated symmetric coefficient matrix are called *positive semi-definite* if

$$q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x} \geq 0 \qquad \text{for all} \qquad \mathbf{x} \in \mathbb{R}^n. \tag{12.8}$$

A positive semi-definite matrix may have *null directions*, meaning non-zero vectors $\mathbf{z}$ such that $q(\mathbf{z}) = \mathbf{z}^T K \mathbf{z} = \mathbf{0}$. Clearly, any nonzero vector $\mathbf{z}$ such that $K \mathbf{z} = \mathbf{0}$ defines a null direction, but there may be others. A positive definite matrix is not allowed to have null directions, and so $\ker K = \{\mathbf{0}\}$. Therefore:

Proposition 12.6. *If K is positive definite, then K is nonsingular.*

The converse, however, is *not* valid; many symmetric, nonsingular matrices fail to be positive definite.

Example 12.7. The matrix $K = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ is positive semi-definite, but not positive definite. Indeed, the associated quadratic form

$$q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x} = x_1^2 - 2 x_1 x_2 + x_2^2 = (x_1 - x_2)^2 \geq 0$$

is a perfect square, and so clearly non-negative. However, the elements of $\ker K$, namely the scalar multiples of the vector $(\, 1, 1 \,)^T$, define null directions: $q(c, c) = 0$.

In a similar fashion, a quadratic form $q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x}$ and its associated symmetric matrix K are called *negative semi-definite* if $q(\mathbf{x}) \leq 0$ for all $\mathbf{x}$ and *negative definite* if $q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$. A quadratic form is called *indefinite* if it is neither positive nor negative semi-definite; equivalently, there exist points $\mathbf{x}_+$ where $q(\mathbf{x}_+) > 0$ and points $\mathbf{x}_-$ where $q(\mathbf{x}_-) < 0$. Details can be found in the exercises.

Gram Matrices

Symmetric matrices whose entries are given by inner products of elements of an inner product space will appear throughout this text. They are named after the nineteenth century Danish mathematician Jorgen Gram — not the metric mass unit!

Definition 12.8. Let A be an $m \times n$ matrix Then the $n \times n$ matrix

$$K = A^T A \tag{12.9}$$

is known as the associated *Gram matrix*.

Example 12.9. If

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{pmatrix}, \qquad \text{then} \qquad K = A^T A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 45 \end{pmatrix}.$$

The resulting matrix is positive definite owing to the following result.

Theorem 12.10. *All Gram matrices are positive semi-definite. The Gram matrix* $K = A^T A$ *is positive definite if and only if* $\ker A = \{0\}$.

Proof: To prove positive (semi-)definiteness of K, we need to examine the associated quadratic form

$$q(\mathbf{x}) = \mathbf{x}^T K \mathbf{x} = \mathbf{x}^T A^T A \mathbf{x} = (A \mathbf{x})^T A \mathbf{x} = \| A \mathbf{x} \|^2 \geq 0,$$

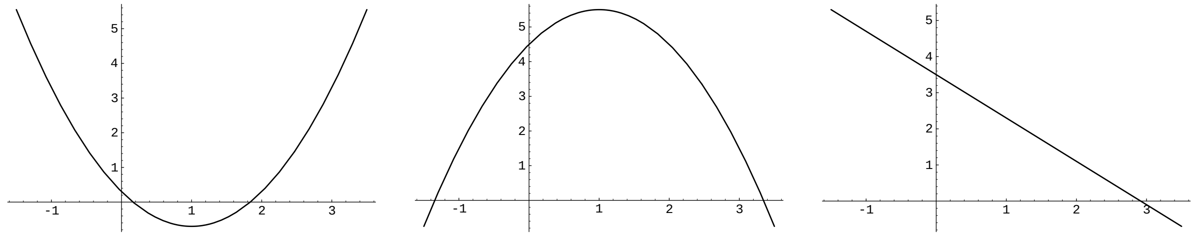
for all $\mathbf{x} \in \mathbb{R}^n$. Moreover, it equals 0 if and only if $A \mathbf{x} = \mathbf{0}$, and so if A has trivial kernel, this requires $\mathbf{x} = \mathbf{0}$, and hence $q(\mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$. Thus, in this case, $q(\mathbf{x})$ and K are positive definite. *Q.E.D.*

More generally, if $C > 0$ is any symmetric, positive definite $m \times m$ matrix, then we define the *weighted Gram matrix*

$$K = A^T C A. \tag{12.10}$$

Theorem 12.10 also holds as stated for weighted Gram matrices. In the majority of applications, $C = \operatorname{diag}(c_1, \ldots, c_m)$ is a diagonal positive definite matrix, which requires it to have strictly positive diagonal entries $c_i > 0$.

Example 12.11. Returning to the situation of Example 12.9, let $C = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ be a diagonal positive definite matrix. Then the corresponding weighted Gram matrix



$a > 0$ $\qquad$ $a < 0$ $\qquad$ $a = 0$

Figure 12.1. Parabolas.

(12.10) is

$$\widetilde{K} = A^T C\, A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 6 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{pmatrix} = \begin{pmatrix} 16 & -21 \\ -21 & 207 \end{pmatrix},$$

which is again positive definite.

12.2. Minimization of Quadratic Functions.

The simplest algebraic equations are linear systems. As such, they must be thoroughly understood before venturing into the far more complicated nonlinear realm. For minimization problems, the starting point is the quadratic function. (Linear functions do not have minima — think of the function $f(x) = \alpha x + \beta$ whose graph is a straight line.) In this section, we shall see how the problem of minimizing a general quadratic function of n variables can be solved by linear algebra techniques.

Let us begin by reviewing the very simplest example — minimizing a scalar quadratic function

$$p(x) = a x^2 + 2 b x + c \tag{12.11}$$

over all possible values of $x \in \mathbb{R}$. If $a > 0$, then the graph of p is a parabola pointing upwards, and so there exists a unique minimum value. If $a < 0$, the parabola points downwards, and there is no minimum (although there is a maximum). If $a = 0$, the graph is a straight line, and there is neither minimum nor maximum — except in the trivial case when $b = 0$ also, and the function $p(x) = c$ is constant, with every x qualifying as a minimum and a maximum. The three nontrivial possibilities are sketched in Figure 12.1.

In the case $a > 0$, the minimum can be found by calculus. The *critical points* of a function, which are candidates for minima (and maxima), are found by setting its derivative to zero. In this case, differentiating, and solving

$$p'(x) = 2 a x + 2 b = 0,$$

we conclude that the only possible minimum value occurs at

$$x^\star = -\,\frac{b}{a}, \qquad \text{where} \qquad p(x^\star) = c - \frac{b^2}{a}. \tag{12.12}$$

Of course, one must check that this critical point is indeed a minimum, and not a maximum or inflection point. The second derivative test will show that $p''(x^\star) = 2a > 0$, and so $x^\star$ is at least a local minimum.

A more instructive approach to this problem — and one that only requires elementary algebra — is to "complete the square". We rewrite

$$p(x) = a\left(x + \frac{b}{a}\right)^2 + \frac{ac - b^2}{a}. \tag{12.13}$$

If $a > 0$, then the first term is always ≥ 0, and, moreover, attains its minimum value 0 only at $x^\star = -b/a$. The second term is constant, and so is unaffected by the value of x. Thus, the global minimum of $p(x)$ is at $x^\star = -b/a$. Moreover, its minimal value equals the constant term, $p(x^\star) = (ac - b^2)/a$, thereby reconfirming and strengthening the calculus result in (12.12).

Now that we have the one-variable case firmly in hand, let us turn our attention to the more substantial problem of minimizing quadratic functions of several variables. Thus, we seek to minimize a (real) *quadratic function*

$$p(\mathbf{x}) = p(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} k_{ij}\, x_i\, x_j - 2\sum_{i=1}^{n} f_i\, x_i + c, \tag{12.14}$$

depending on n variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$. The coefficients k_{ij}, f_i and c are all assumed to be real. Moreover, we can assume, without loss of generality, that the coefficients of the quadratic terms are symmetric: $k_{ij} = k_{ji}$. Note that $p(\mathbf{x})$ is more general than a quadratic form (12.2) in that it also contains linear and constant terms. We seek a global minimum, and so the variables $\mathbf{x}$ are allowed to vary over all of $\mathbb{R}^n$.

Let us begin by rewriting the quadratic function (12.14) in a more compact matrix notation:

$$p(\mathbf{x}) = \mathbf{x}^T K\, \mathbf{x} - 2\, \mathbf{x}^T \mathbf{f} + c, \tag{12.15}$$

in which $K = (k_{ij})$ is a symmetric $n \times n$ matrix, $\mathbf{f}$ is a constant vector, and c is a constant scalar. We first note that in the simple scalar case (12.11), we needed to impose the condition that the quadratic coefficient a is *positive* in order to obtain a (unique) minimum. The corresponding condition for the multivariable case is that the quadratic coefficient matrix K be *positive definite*. This key assumption enables us to establish a general minimization criterion.

Theorem 12.12. *If K is a symmetric, positive definite matrix, then the quadratic function* (12.15) *has a unique minimizer, which is the solution to the linear system*

$$K\,\mathbf{x} = \mathbf{f}, \qquad \textit{namely} \qquad \mathbf{x}^\star = K^{-1}\mathbf{f}. \tag{12.16}$$

The minimum value of $p(\mathbf{x})$ is equal to any of the following expressions:

$$p(\mathbf{x}^\star) = p(K^{-1}\mathbf{f}) = c - \mathbf{f}^T K^{-1}\mathbf{f} = c - \mathbf{f}^T \mathbf{x}^\star = c - (\mathbf{x}^\star)^T K\, \mathbf{x}^\star. \tag{12.17}$$

Proof: First recall that, by Proposition 12.6, positive definiteness implies that K is a nonsingular matrix, and hence the linear system (12.16) does have a unique solution $\mathbf{x}^\star = K^{-1}\mathbf{f}$. Then, for any $\mathbf{x} \in \mathbb{R}^n$, we can write

$$\begin{aligned} p(\mathbf{x}) &= \mathbf{x}^T K\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{f} + c = \mathbf{x}^T K\,\mathbf{x} - 2\,\mathbf{x}^T K\,\mathbf{x}^\star + c \\ &= (\mathbf{x}-\mathbf{x}^\star)^T K(\mathbf{x}-\mathbf{x}^\star) + \left[\, c - (\mathbf{x}^\star)^T K\,\mathbf{x}^\star \,\right], \end{aligned} \tag{12.18}$$

where we used the symmetry of $K = K^T$ to identify $\mathbf{x}^T K\,\mathbf{x}^\star = (\mathbf{x}^\star)^T K\,\mathbf{x}$. The first term in the final expression has the form $\mathbf{y}^T K\,\mathbf{y}$, where $\mathbf{y} = \mathbf{x} - \mathbf{x}^\star$. Since we assumed that K is positive definite, we know that $\mathbf{y}^T K\,\mathbf{y} > 0$ for all $\mathbf{y} \neq \mathbf{0}$. Thus, the first term achieves its minimum value, namely 0, if and only if $\mathbf{0} = \mathbf{y} = \mathbf{x} - \mathbf{x}^\star$. Moreover, since $\mathbf{x}^\star$ is fixed, the second term does not depend on $\mathbf{x}$. Therefore, the minimum of $p(\mathbf{x})$ occurs at $\mathbf{x} = \mathbf{x}^\star$ and its minimum value $p(\mathbf{x}^\star)$ is equal to the constant term. The alternative expressions in (12.17) follow from simple substitutions. *Q.E.D.*

Example 12.13. Consider the problem of minimizing the quadratic function

$$p(x_1, x_2) = 4\,x_1^2 - 2\,x_1\,x_2 + 3\,x_2^2 + 3\,x_1 - 2\,x_2 + 1$$

over all (real) x_1, x_2. We first write the function in our matrix form (12.15), so

$$p(x_1,x_2) = (\,x_1 \;\; x_2\,) \begin{pmatrix} 4 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2\,(\,x_1 \;\; x_2\,) \begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix} + 1,$$

whereby

$$K = \begin{pmatrix} 4 & -1 \\ -1 & 3 \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix}. \tag{12.19}$$

(Pay attention to the overall factor of -2 in front of the linear terms.) According to Theorem 12.12, to find the minimum we must solve the linear system

$$\begin{pmatrix} 4 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix}. \tag{12.20}$$

Applying the usual Gaussian Elimination algorithm, only one row operation is required to place the coefficient matrix in upper triangular form:

$$\left(\begin{array}{cc|c} 4 & -1 & -\frac{3}{2} \\ -1 & 3 & 1 \end{array}\right) \longmapsto \left(\begin{array}{cc|c} 4 & -1 & -\frac{3}{2} \\ 0 & \frac{11}{4} & \frac{5}{8} \end{array}\right).$$

The coefficient matrix is regular as no row interchanges were required, and its two pivots, namely $4, \frac{11}{4}$, are both positive. Thus, by Theorem 12.4, $K > 0$ and hence $p(x_1, x_2)$ really does have a minimum, obtained by applying Back Substitution to the reduced system:

$$\mathbf{x}^\star = \begin{pmatrix} x_1^\star \\ x_2^\star \end{pmatrix} = \begin{pmatrix} -\frac{7}{22} \\ \frac{5}{22} \end{pmatrix} \approx \begin{pmatrix} -.31818 \\ .22727 \end{pmatrix}. \tag{12.21}$$

The quickest way to compute the minimal value

$$p(\mathbf{x}^\star) = p\left(-\tfrac{7}{22}, \tfrac{5}{22}\right) = \tfrac{13}{44} \approx .29546$$

is to use the second formula in (12.17).

It is instructive to compare the algebraic solution method with the minimization procedure you learned in multi-variable calculus, cf. [**2**, **36**]. The *critical points* of $p(x_1, x_2)$ are found by setting both partial derivatives equal to zero:

$$\frac{\partial p}{\partial x_1} = 8x_1 - 2x_2 + 3 = 0, \qquad \frac{\partial p}{\partial x_2} = -2x_1 + 6x_2 - 2 = 0.$$

If we divide by an overall factor of 2, these are precisely the *same* linear equations we already constructed in (12.20). Thus, not surprisingly, the calculus approach leads to the same minimizer (12.21). To check whether $\mathbf{x}^\star$ is a (local) minimum, we need to apply the second derivative test. In the case of a function of several variables, this requires analyzing the *Hessian matrix*, which is the symmetric matrix of second order partial derivatives

$$H = \begin{pmatrix} \dfrac{\partial^2 p}{\partial x_1^2} & \dfrac{\partial^2 p}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 p}{\partial x_1 \partial x_2} & \dfrac{\partial^2 p}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 8 & -2 \\ -2 & 6 \end{pmatrix} = 2K,$$

which is exactly twice the quadratic coefficient matrix (12.19). If the Hessian matrix is positive definite — which we already know in this case — then the critical point is indeed a (local) minimum.

Thus, the calculus and algebraic approaches to this minimization problem lead, as they must, to identical results. However, the algebraic method is *more* powerful, because it immediately produces the *unique*, *global* minimum, whereas, barring additional work, calculus can only guarantee that the critical point is a local minimum. Moreover, the proof of the calculus local minimization criterion — that the Hessian matrix be positive definite at the critical point — relies, in fact, on the algebraic solution to the quadratic minimization problem! In summary: minimization of quadratic functions is a problem in linear algebra, while minimizing more complicated functions requires the full force of multivariable calculus.

The most efficient method for producing a minimum of a quadratic function $p(\mathbf{x})$ on $\mathbb{R}^n$, then, is to first write out the symmetric coefficient matrix K and the vector $\mathbf{f}$. Solving the system $K\mathbf{x} = \mathbf{f}$ will produce the minimizer $\mathbf{x}^\star$ *provided* $K > 0$ — which should be checked during the course of the procedure using the criteria of Theorem 12.4, that is, making sure that no row interchanges are used and all the pivots are positive.

Example 12.14. Let us minimize the quadratic function

$$p(x, y, z) = x^2 + 2xy + xz + 2y^2 + yz + 2z^2 + 6y - 7z + 5.$$

This has the matrix form (12.15) with

$$K = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 2 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} 0 \\ -3 \\ \frac{7}{2} \end{pmatrix}, \qquad c = 5.$$

Gaussian Elimination produces the $L\,D\,L^T$ factorization

$$K = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{7}{4} \end{pmatrix} \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The pivots, i.e., the diagonal entries of D, are all positive, and hence K is positive definite. Theorem 12.12 then guarantees that $p(x, y, z)$ has a unique minimizer, which is found by solving the linear system $K\mathbf{x} = \mathbf{f}$. The solution is then quickly obtained by forward and back substitution:

$$x^\star = 2, \qquad y^\star = -3, \qquad z^\star = 2, \qquad \text{with} \qquad p(x^\star, y^\star, z^\star) = p(2, -3, 2) = -11.$$

Theorem 12.12 solves the general quadratic minimization problem when the quadratic coefficient matrix is positive definite. If K is not positive definite, then the quadratic function (12.15) does not have a minimum, apart from one exceptional situation.

Theorem 12.15. *If K is positive definite, then the quadratic function $p(\mathbf{x}) = \mathbf{x}^T K\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{f} + c$ has a unique global minimizer $\mathbf{x}^\star$ satisfying $K\,\mathbf{x}^\star = \mathbf{f}$. If K is only positive semi-definite, and $\mathbf{f} \in \operatorname{rng} K$, then every solution to the linear system $K\,\mathbf{x}^\star = \mathbf{f}$ is a global minimum of $p(\mathbf{x})$, but the minimum is not unique since $p(\mathbf{x}^\star + \mathbf{z}) = p(\mathbf{x}^\star)$ for any null vector $\mathbf{z} \in \ker K$. In all other cases, $p(\mathbf{x})$ has no global minimum.*

