The spectral theorem and local quadratic approximations

Mark S. Gockenbach

1 Introduction

If $f : \mathrm{R}^n \to \mathrm{R}$ is twice differentiable at $a \in \mathrm{R}^n$, then f can be approximated near $x = a$ by a quadratic:

$$f(x) \doteq f(a) + \nabla f(a) \cdot (x-a) + \frac{1}{2}(x-a) \cdot \nabla^2 f(a)(x-a).$$

Written out in coordinates, this approximation is rather complicated:

$$f(x) \doteq f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)(x_i - a_i)(x_j - a_j). \tag{1}$$

However, $\nabla^2 f(a)$ is a symmetric matrix, and its *spectral decomposition* (that is, its expression in terms of eigenvalues and eigenvectors) makes it possible to considerably simplify (1).

2 The spectral theorem for symmetric matrices

Symmetric matrices have many special properties, the most important of which are expressed in the following theorem:

Theorem 2.1 *Suppose $A \in \mathrm{R}^{n \times n}$ is symmetric. Then*

1. *every eigenvalue λ of A is a real number and there exists a (real) eigenvector $u \in \mathrm{R}^n$ corresponding to λ: $Au = \lambda u$;*
2. *eigenvectors corresponding to distinct eigenvalues are necessarily orthogonal:*

$$Au^{(1)} = \lambda_1 u^{(1)},\ Au^{(2)} = \lambda_2 u^{(2)},\ \lambda_1 \neq \lambda_2 \ \Rightarrow\ u^{(1)} \cdot u^{(2)} = 0.$$

3. *there exists a diagonal matrix $D \in \mathrm{R}^{n \times n}$ and an orthogonal matrix $U \in \mathrm{R}^{n \times n}$ such that $A = UDU^T$. The diagonal entries of D are the eigenvalues of A and the columns of U are the corresponding eigenvectors:*

$$D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n),\ U = [u^{(1)}|u^{(2)}|\ldots|u^{(n)}],\ Au^{(i)} = \lambda_i u^{(i)},\ i = 1, 2, \ldots, n.$$

An orthogonal matrix U satisfies, by definition, $U^T = U^{-1}$, which means that the columns of U are orthonormal (that is, any two of them are orthogonal and each has norm one). The expression $A = UDU^T$ of a symmetric matrix in terms of its eigenvalues and eigenvectors is referred to as the *spectral decomposition* of A.

The spectral theorem implies that there is a change of variables which transforms A into a diagonal matrix. Before explaining this change of variables, I will show why it is important. The reader will recall that every quadratic function in the n variables $x_1, x_2, \ldots, x_n$ can be expressed in the form

$$q(x) = x \cdot Hx = \sum_{i=1}^{n} \sum_{j=1}^{n} H_{ij} x_i x_j.$$

The formula for $q(x)$ involves n^2 terms, and the variables are typically coupled. However, if H happens to be a diagonal matrix, then the formula for $q(x)$ simplifies considerably:

$$q(x) = \sum_{i=1}^{n} H_{ii} x_i^2.$$

Such a quadratic is easy to understand: In each coordinate direction x_i, the graph is a parabola, opening upward if $H_{ii} > 0$ and opening downward if $H_{ii} < 0$. There is also the degenerate case $H_{ii} = 0$, in which case q is constant with respect to x_i and the graph in that direction is a horizontal line.

Therefore, in two variables (the only case that can be visualized), a quadratic function defined by $H = \operatorname{diag}(\lambda_1, \lambda_2)$ has six possible shapes, corresponding to the following cases:

1. $\lambda_1 > 0, \lambda_2 > 0$;
2. $\lambda_1 < 0, \lambda_2 < 0$;
3. $\lambda_1 > 0, \lambda_2 < 0$ or $\lambda_1 < 0, \lambda_2 > 0$;
4. $\lambda_1 > 0, \lambda_2 = 0$ or $\lambda_1 = 0, \lambda_2 > 0$;
5. $\lambda_1 < 0, \lambda_2 = 0$ or $\lambda_1 = 0, \lambda_2 < 0$;
6. $\lambda_1 = 0, \lambda_2 = 0$.

Four of the possibilities are graphed in Figure 1.

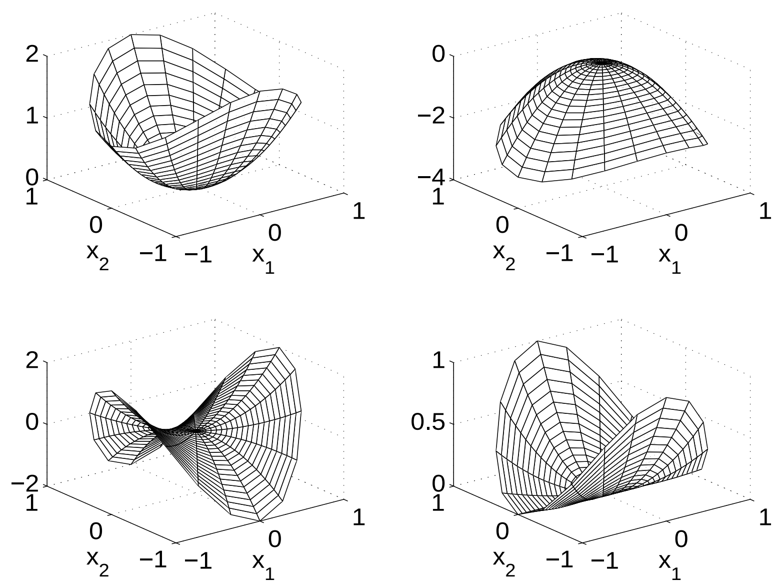

Figure 1: The graphs of four quadratic functions: two positive eigenvalues (upper left), two negative eigenvalues (upper right), one positive and one negative eigenvalue (lower left), one positive and one zero eigenvalue (lower right).

Now I will explain the change of variables that *diagonalizes* a symmetric matrix. A vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

is implicitly expressed in terms of the *standard basis* $e^{(1)}, e^{(2)}, \ldots, e^{(n)}$:

$$x = \sum_{i=1}^{n} x_i e^{(i)},$$

where

$$e^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; e^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; \ldots.$$

If $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$ is an orthonormal set, then it is an alternate basis: Every $x \in \mathrm{R}^n$ can be expressed as

$$x = \sum_{i=1}^{n} \alpha_i u^{(i)}.$$

Moreover, the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ are easy to compute:

$$\alpha_i = u^{(i)} \cdot x, \; i = 1, 2, \ldots, n.$$

When the orthonormal basis forms a matrix $U = [u^{(1)}|u^{(2)}|\ldots|u^{(n)}]$, then the computation of the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ takes for the form of a matrix-vector product:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} u^{(1)} \cdot x \\ u^{(2)} \cdot x \\ \vdots \\ u^{(n)} \cdot x \end{bmatrix} = U^T x.$$

The key point here is that the numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ can be thought of as new variables representing the vector x. Specifically, $x_1, x_2, \ldots, x_n$ represent x in the standard basis $\{e^{(1)}, e^{(2)}, \ldots, e^{(n)}\}$, while $\alpha_1, \alpha_2, \ldots, \alpha_n$ represent x in the alternate basis $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$.

I now digress to remind the reader of the following fundamental property of matrices, vectors, and the dot product: If $A \in \mathrm{R}^{m \times n}$, then

$$y \cdot Ax = (A^T y) \cdot x \;\text{ for all } x \in \mathrm{R}^n, y \in \mathrm{R}^m.$$

This is really the reason that the transpose of a matrix is important.

Assuming $H \in \mathrm{R}^{n \times n}$ is symmetric, it has a spectral decomposition $H = UDU^T$. Therefore,

$$x \cdot Hx = x \cdot UDU^T x = (U^T x) \cdot D(U^T x) = \sum_{i=1}^{n} \lambda_i \alpha_i^2,$$

where I have applied the change of variables $\alpha = U^T x$. Therefore, the quadratic $q(x) = x \cdot Hx$ is a simple decoupled quadratic when expressed in terms of the alternate basis $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$. Since every symmetric matrix has a spectral decomposition, this means that *every* quadratic function $q(x) = x \cdot Hx$ can be expressed as a simple decoupled quadratic, provided the correct coordinate system is chosen. In particular, this shows that the graph of every quadratic in two variables looks like one of the graphs in Figure 1 (or like one of the two other possibilities not illustrated in that figure), possibly rotated from the standard coordinates.

Example 2.2 *Define* $q : \mathrm{R}^2 \to \mathrm{R}$ *by*

$$q(x) = x_1^2 + 6x_1 x_2 + x_2^2.$$

Then $q(x) = x \cdot Hx$, *where*

$$H = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}.$$

The spectral decomposition of H *is* $H = UDU^T$, *where*

$$D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}, \ U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

The vectors

$$u^{(1)} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \ u^{(2)} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

define the coordinate system illustrated in Figure 2. The graph of q, *which is shown in Figure 3, is*

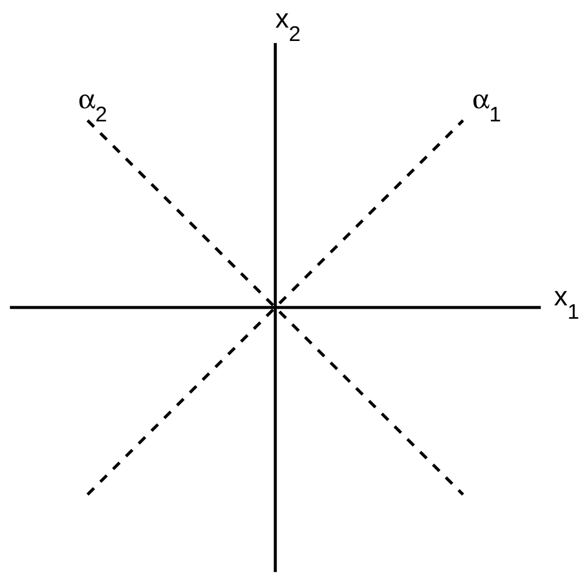

Figure 2: Standard coordinates and a rotated coordinate system.

now predictable: It curves up in the direction of $u^{(1)}$ *and down in the direction of* $u^{(2)}$.

3 Understanding stationary points

If x^* is a stationary point of $f : \mathrm{R}^n \to \mathrm{R}$, then

$$f(x) \doteq f(x^*) + \frac{1}{2}(x - x^*) \cdot \nabla^2 f(x^*)(x - x^*) \ (x \text{ near } x^*).$$

Changing variables to $y = x - x^*$ and writing $H = \nabla^2 f(x^*)$ yields

$$f(x^* + y) \doteq f(x^*) + \frac{1}{2} y \cdot Hy \ (y \text{ near } 0).$$

Since H is symmetric and $f(x^*)$ is a constant, this shows that, near x^*, the graph of f looks like a quadratic, and I can now distinguish several possibilities:

1. All the eigenvalues of H are positive, so the graph of f "curves upward" in every direction, and x^* is a strict local minimizer. I can also say that f looks like a convex quadratic near x^*.

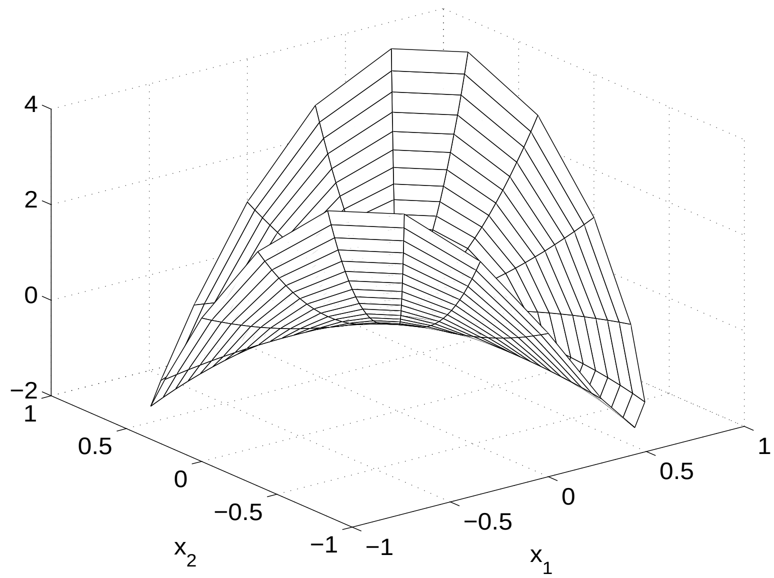

Figure 3: The function $q(x) = x_1^2 + 6x_1x_2 + x_2^2$.

2. All the eigenvalues of H are negative, so the graph of f "curves downward" in every direction, and x^* is a strict local maximizer. Near x^*, f can be approximated by a concave quadratic.

3. Some eigenvalues of H are positive and others are negative. In this case, f is convex in some directions and concave in others. The point x^* is called a *saddle point* of f. (The function graphed in Figure 3 has a saddle point at the origin.)

4. One or more eigenvalues of H is zero. In this case, f is locally constant in the corresponding eigendirections, and $\nabla^2 f(x^*)$ may not reveal whether f has a local maximum, a local minimum, or a saddle point.