Positive Definite Matrices 1

Dhruv Kohli

Department of Mathematics
Indian Institute of Technology, Guwahati

Outline

Motivation

- Signs of eigenvalues are crucial.
  - Stability of differential equations governed by sign of λ. Example: $e^{\lambda t}$ is unstable for $\lambda > 0$.
- Finding Minima of a function is an extremely important problem. For single variable functions the tests are $f'(a) = 0$ and $f''(a) > 0$. Analogous tests required for multivaraible functions.
- When a multiple of one row is subtracted from another, the row space, nullspace, rank and determinant, all remain the same. The eigenvalues are unchanged with similarity transformation. What are the elementary operations and invariants for $x^T Ax$?

Tests for Minima in 2-variable functions

- We want to check whether $F(x,y)$ has minima at (a,b).
- **Step 1**: Check whether (a,b) is a stationary point:

$$F_x(a,b) = 0 \text{ and } F_y(a,b) = 0$$

- If yes, then $z = F(a,b)$ will be tangential to $z = F(x,y)$.
- **Step 2**: Check whether the neighbourhood of (a,b) is above $F(a,b)$ or below or both. Now, we construct a test based on second order terms. First, note that the **quadratic part of** F is (assuming $F_{xy} = F_{yx}$):

$$g(x,y) = \frac{x^2}{2}F_{xx}(a,b) + xyF_{xy}(a,b) + \frac{y^2}{2}F_{yy}(a,b)$$

- $g(x,y)$ behaves near $(0,0)$ in the same manner as $F(x,y)$ near (a,b).[1] Note that $g_x(0,0) = 0$ and $g_y(0,0) = 0$.

[1] Ignoring singular case when each second order term is zero, in which case, higher order terms are drawn into the problem.

Tests for Minima in 2-variable functions contd.

- Consider $g(x,y) = ax^2 + 2bxy + cy^2$. Note, $g(0,0) = 0$, $g_x(0,0) = 0$ and $g_y(0,0) = 0$. So, $g(x,y)$ has a stationary point at $(0,0)$. When $g(x,y)$ is strictly positive at points other than $(0,0)$, then $g(x,y)$ is called **positive definite** and it has a minima at $(0,0)$.
- **What conditions on a, b and c ensure that $g(x,y)$ is positive definite?**

$$g(1,0) > 0 \Rightarrow a > 0$$
$$g(0,1) > 0 \Rightarrow c > 0$$

- A large cross term can still pull graph below zero. Example: $h(x,y) = x^2 - 10xy + y^2$ has $a = c = 1 > 0$ but $h(1,1) = -8 < 0$. Also, the sign of b is of no importance. Example: $h(x,y) = 2x^2 + 4xy + y^2$ has $a, b, c > 0$ but $h(1,-1) = -1 < 0$. **It is the size of b compared to a and c that must be controlled.**

Tests for Minima in 2-variable functions contd.

- Express $g(x, y)$ using squares:

$$g(x, y) = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2 \tag{1}$$

- So, the third necessary condition is:

$$g(x, y) > 0 \Rightarrow ac > b^2$$

Result 1

g(x,y) $= ax^2 + 2bxy + cy^2$ is positive definite $\iff$ $a > 0$ and $ac > b^2$.[2] The surface $z = g(x, y)$ will be a bowl opeining up. Any $F(x, y)$ has a minimum at a point (a, b) if

$$F_x(a, b) = 0, \;\; F_y(a, b) = 0$$

$$F_{xx}(a, b) > 0 \;\; \text{and} \;\; F_{xx}(a, b)F_{yy}(a, b) > F_{xy}(a, b)^2$$

[2] $a > 0$ and $ac > b^2 \Rightarrow c > 0$

Other tests in 2-variable functions

- **Maximum**: $g(x, y)$ has a maximum whenever $-g(x, y)$ has a minimum. So, the necessary and sufficient conditions for a maximum can be derived by reversing signs of a, b and c in conditions for a minimum: $a < 0$ and $ac > b^2$.
- **Singular case $\mathbf{ac = b^2}$**: Equation 1 reduces to $g(x, y) = a\left(x + \frac{b}{a}y\right)^2$ which is **positive semidefinite** if $a > 0$ and **negative semidefinite** if $a < 0$. The surface of $z = g(x, y)$ will be a valley (not a bowl) running along $x + \frac{b}{a}y = 0$. Example: $g(x, y) = (x + y)^2$.
- **Saddle point $\mathbf{ac < b^2}$** The neighborhood of $g(x, y)$ at $(0, 0)$ will have points both above and below zero. Therefore, $(0, 0)$ is neither a maximum nor a minimum and is called a **saddle point** and $g(x, y)$ is **indefinite**. Example: $h(x, y) = 2xy$, $h(x, y) = x^2 - y^2$.

Higher Dimensions

- $g(x,y) = ax^2 + 2bxy + cy^2$ can be expressed in matrix notation as:

$$ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

  - Note that a, b and c are second order terms.
- Similarly, for any symmetric matrix A, the product $x^T Ax$ is a pure quadratic form $P(x_1, x_2, \ldots, x_n)$:

$$\begin{bmatrix} x_1 & x_2 & . & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & . & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

  - Note: $P(\mathbf{0}) = 0$ and $\mathbf{0}$ is a stationary point. A contains second order terms where $P_{x_i x_j}(\mathbf{0}) = P_{x_j x_i}(\mathbf{0})$, implying, A is symmetric. **P has a minimum at 0 when the pure quadratic form $\mathbf{x}^{\mathbf{T}}\mathbf{Ax}$ is positive definite.**

Tests for Positive Definiteness

Result 2

Each of the following tests is a necessary and sufficient condition for the real symmetric matrix A to be **positive definite**:

1. $x^T Ax > 0$ for all nonzero real vectors x.
2. All the eigenvalues of A satisfy $\lambda_i > 0$.
3. All the upper left submatrices A_k have positive determinants.
4. All the pivots (without row exchanges) satisfy $d_k > 0$.

Before proving this result, we check it on $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

Tests for Positive Definiteness - 2×2 case

- Consider $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \xrightarrow{\text{after elimination}} \begin{bmatrix} a & b \\ 0 & c - \frac{b^2}{a} \end{bmatrix}$.
- From result 1, A is positive definite if and only if $a > 0$ and $ac > b^2$ $(\Rightarrow c > 0)$.

- $\mathbf{1} \iff \mathbf{2}$:

$$\begin{matrix} \lambda_1 + \lambda_2 = tr(A) = a + c > 0 \\ \lambda_1\lambda_2 = det(A) = ac - b^2 > 0 \end{matrix} \iff \lambda_1 > 0,\ \lambda_2 > 0$$

- $\mathbf{1} \iff \mathbf{3}$:

$$\det(A_1) = a > 0,\ \det(A_2) = ac - b^2 > 0$$

- $\mathbf{1} \iff \mathbf{4}$:

$$d_1 = a > 0,\ d_2 = c - \frac{b^2}{a} > 0$$

- Note: $\mathbf{3} \Rightarrow \mathbf{4}$ since $d_k = \frac{det(A_k)}{det(A_{k-1})}$ where $det(A_0) = 1$.

Tests for Positive Definiteness - Proof Hints

$1 \iff 2$

$x_i^T A x_i = \lambda_i x_i^T x_i \Rightarrow \lambda_i = \frac{x_i^T A x_i}{x_i^T x_i}$.

$A = A^T \Rightarrow A$ has full set of orthonormal evecs $x_1, x_2, \ldots, x_n$.

$$x = \sum_{i=1}^{n} c_i x_i \Rightarrow Ax = \sum_{i=1}^{n} c_i \lambda_i x_i \Rightarrow x^T A x = \sum_{i=1}^{n} c_i^2 \lambda_i > 0$$

$1 \Rightarrow 3$

Take $x = (x_k, 0)$ where last $n-k$ terms are 0.

$x^T A x > 0 \Rightarrow x_k^T A_k x_k > 0 \Rightarrow \lambda_{k_i} > 0 \Rightarrow det(A_k) = \prod \lambda_{k_i} > 0$

$3 \Rightarrow 4,\ 4 \Rightarrow 1$

$d_k = det(A_k)/det(A_{k-1}),\ det(A_0) = 1$.

$A = A^T \Rightarrow A = LDL^T \Rightarrow x^T A x = y^T D y = \sum_i d_i y_i^2,\ y = L^T x$

Remarks - Result 2

- Result 2 connects: Eigenvalues, Determinants and Pivots. Each test is enough by itself.
- $4 \implies 1$ shows that elimination and completing the square are actually the same. Example:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = LDL^T,$$

$$L^T = \begin{bmatrix} 1 & \frac{-1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix}$$

$$y = L^T x = \left[x_1 - \tfrac{x_2}{2}, x_2 - \tfrac{2}{3}x_3, x_3\right]^T$$

$$x^T A x = y^T D y = 2\left(x_1 - \frac{x_2}{2}\right)^2 + \frac{3}{2}\left(x_2 - \frac{2}{3}x_3\right)^2 + \frac{4}{3}x_3^2$$

Result 3

The symmetric matrix A is positive definite $\iff$ there exists a matrix R with independent columns such that $A = R^T R$.

Proof - Hints:

$(\Leftarrow)$ $x^T R^T R x = (Rx)^T (Rx) > 0$

$(\Rightarrow)$ **Elimination**: $A = LDL^T = R^T R$ where $R = \sqrt{D} L^T$.
Since A's pivots $d_k > 0 \Rightarrow \text{rank}(A) = n$ and therefore, columns of R are independent.

$(\Rightarrow)$ **Eigenvalues**: $A = Q \Lambda Q^T = R^T R$ where $R = \sqrt{\Lambda} Q^T$.
Since $\Lambda_{ii} > 0$ and Q's columns are orthonormal, therefore, columns of R are independent.

$(\Rightarrow)$ **Symmetric Positive Def. Square Root**: $A = Q\sqrt{\Lambda} Q^T$.

$(\Rightarrow)$ **Many Choices**: Let Q be a matrix with orthonormal columns, then $(QR)^T (QR) = R^T Q^T Q R = R^T R$.

Tests for Positive Semidefiniteness

Result 3

Each of the following tests is a necessary and sufficient condition for the real symmetric matrix A to be **positive semidefinite**:

1. $x^T A x \geq 0$ for all nonzero real vectors x.
2. All the eigenvalues of A satisfy $\lambda_i \geq 0$.
3. All the principal submatrices have nonnegative determinants. All principal minors are nonnegative.
4. All the pivots satisfy $d_k \geq 0$.
5. $\exists R$ possibly with dependent columns such that $A = R^T R$.

Remark on proof: Add ϵI to A so that $A + \epsilon I$ is positive definite. Then let ϵ approach 0. Since determinants and eigenvalues depend continuously on ϵ, they will be positive until the very last moment. At $\epsilon = 0$, they must still be nonnegative.

What if A is unsymmetric?

- A reasonable definition of unsymmetric positive definite matrices is: $\frac{1}{2}(A + A^H)$ should be positive definite. That guarantees that the real parts of eigenvalues of A are positive but the converse is not true.

$$x^H A x = \lambda x^H x, \; x^H A^H x = \bar{\lambda} x^H x$$

$$\Rightarrow Re(\lambda) = \frac{x^H (A + A^H) x}{2 x^H x} > 0$$

  - Counterexample for converse: $A = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ has $\lambda > 0$, but $\frac{1}{2}(A + A^T) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is indefinite.

Ellipsoids in n Dimensions

- When A is positive definite then $x^T Ax = 1$ defines a curved figure called **ellipsoid** in n dimensions.
- When $A = I$ then it is a sphere in n dimensions.
- When A is diagonal then A is ellipsoid with axes lined up with coordinate axes.
- When $A \neq I$ then the axes point towards eigenvectors of A with major axis pointing toward eigenvector corresponding to least eigenvalue.

$$x^T Ax = 1 \Rightarrow x^T Q\Lambda Q^T x = 1 \Rightarrow y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 = 1$$

- The resulting ellipsoid has length $\frac{1}{\sqrt{\lambda_i}}$ from origin along the eigenvector Q_i. The change from x to $y = Q^T x$ rotates the axes of the space to match the axes of the ellipsoid.

The Law of Inertia

- **Congruence transformation**: $A \rightarrow C^T AC$ for some nonsingular C.

Result 4

$C^T AC$ has the same number of positive eigenvalues, negative eigenvalues and zero eigenvalues.

Proof Try to prove in some other way.

The Law of Inertia - Application

Result 5

For any symmetric matrix A, the signs of the pivots agree with the signs of eigenvalues. The eigenvalue matrix Λ and the pivots matrix D has the same number of positive entries, negative entries and zero entries.

Proof - Hints: Use $A = LDL^T$ (without row exchanges) and result 4 and conclude that signs of eigenvalues of A and D (for which eigenvalues are pivots) agree.

The Generalized Eigenvalue Problem

- $\mathbf{Ax} = \lambda \mathbf{Mx}$ is the **generalized eigenvalue problem** which involves two matrices instead of one as in $Ax = \lambda x$.
- To solve this problem, we proceed in the same way as the standard eigenvalue problem, i.e., $det(A - \lambda M) = 0$.
- If M is symmetric (then A is also symmetric) and is positive definite, then simplification of the generalized eigenvalue problem to standard eigenvalue problem is possible, since M can be decomposed into $R^T R$:

$$Ax = \lambda Mx = \lambda R^T R x \xrightarrow{y=Rx} AR^{-1}y = \lambda R^T y$$

$$C^T A C y = \lambda y, \;\; C = R^{-1}$$

- The eigenvalues are same as for the original problem and the eigenvectors are related by $y_j = Rx_j$.

The Generalized Eigenvalue Problem contd.

Result 6

When $A = A^T$ and M is symmetric and positive definite, then, $Ax = \lambda Mx$ reduces to $C^TACy = \lambda y$ where $C = R^{-1}$ and $M = R^TR$, leading to the following properties:

1. Eigenvalues for $Ax = \lambda Mx$ are real $\because$ C^TAC is symmetric.
2. The λ's have same sign as sign of A by the law of inertia.
3. C^TAC has orthogonal eigenvectors y_j. So, the eigenvectors of $Ax = \lambda Mx$ have **M-orthogonality**,
$$y_i^Ty_j = 0 \iff x_i^TR^TRx_j = 0 \iff x_i^TMx_j = 0$$
4. A and M are simultaneously diagonalized.
$$C^TAC = Q\Lambda Q^T \Rightarrow S^TAS = \Lambda, \ S = CQ$$
$$S^TMS = (R^{-1}Q)^TR^TR(R^{-1}Q) = Q^TQ = I$$
This is congruence (S^T) not similarity transformation (S^{-1}).

Bibliography

Gilbert Strang. *Linear algebra and its applications*. Belmont, CA: Thomson, Brooks/Cole, 2006.