MATH - UA 140 - Linear Algebra

Lecture 24: Positive Definite Matrices

I] Positive Definite Matrices

1) Definition

Let A be a symmetric matrix. If all its eigenvalues are strictly positive, $\lambda > 0$, we say that A is positive definite.

2) Example

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is positive definite. Its only eigenvalue is 1.

$\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$ is symmetric but not positive definite: the matrix is singular, so 0 is an eigenvalue.

In general, it seems like determining whether a matrix is positive definite can be tedious, as it seems to require the calculation of its n eigenvalues, through the calculation of a determinant. The first point of this lecture is to give faster methods for deciding whether a matrix is positive definite or not.
In the second part of the lecture, we will show an application where the importance of positive definite matrices is highlighted.

II) Tests for positive definite matrices

1) Test with pivots

In lecture 23, we saw that the pivots and eigenvalues of a symmetric matrix have the same signs. Therefore, a matrix is positive definite if and only if all n pivots are positive.

Example: The pivots of $\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$ are 3 and 2 so the matrix is positive definite.

2) Test with determinants

In Lecture 19, we saw how to compute the pivots of a matrix using the successive upper left determinants. The k^{th} pivot is given by

$$d_k = \frac{\det A_k}{\det A_{k-1}}$$

where A_k is the $k \times k$ matrix in the left-hand corner of A. This result gives us another criterion for a matrix to be positive definite:

A matrix is positive definite if and only if all n upper left determinants are positive.

Example: • $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

$A_1 = |2| > 0$; $A_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0$; $A_3 = 2A_2 + \begin{vmatrix} -1 & -1 \\ 0 & 2 \end{vmatrix} = 4 > 0$

so A is positive definite.

• For 2x2 matrices, the criterion has a simple form:

If $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, we need $|a| = a > 0$ and $\begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 > 0$

This makes sense: the product of the two eigenvalues
$\lambda_1 \lambda_2 = \det A = ac - b^2$, so if $\lambda_1 > 0$, $\lambda_2 > 0$, necessarily
$ac - b^2 > 0$

Furthermore, if $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_1 + \lambda_2 > 0$, so $a + c > 0$, and
since $ac > b^2$, a and c have the same sign, so a and
c must be positive, i.e $a > 0$.

3) "Energy" criterion

Consider a symmetric, positive definite matrix A.
For any eigenvector $\vec{x}$, $A\vec{x} = \lambda \vec{x}$, with $\lambda > 0$
Multiplying the equality with $\vec{x}^T$, $\vec{x}^T A \vec{x} = \lambda \vec{x}^T \vec{x} = \lambda \|\vec{x}\|^2 > 0$
So for any eigenvector, $\vec{x}^T A \vec{x} > 0$

Now, since A is symmetric, any vector $\vec{x}$ can be written in an orthonormal basis of eigenvectors $(\vec{x}_1, \dots, \vec{x}_n)$:

$\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \dots + c_n\vec{x}_n$

Then $A\vec{x} = c_1\lambda_1\vec{x}_1 + c_2\lambda_2\vec{x}_2 + \dots + c_n\lambda_n\vec{x}_n$

$$\vec{x}^T A\vec{x} = (c_1\vec{x}_1^T + c_2\vec{x}_2^T + \dots + c_n\vec{x}_n^T)(c_1\lambda_1\vec{x}_1 + c_2\lambda_2\vec{x}_2 + \dots + c_n\lambda_n\vec{x}_n)$$
$$= c_1^2\lambda_1\|\vec{x}_1\|^2 + c_2^2\lambda_2\|\vec{x}_2\|^2 + \dots + c_n^2\lambda_n\|\vec{x}_n\|^2 > 0$$

In other words, A is positive definite if $\vec{x}^T A\vec{x} > 0$ for every nonzero vector $\vec{x}$.

Some mathematicians view this criterion as the most fundamental definition of positive definite matrices.

In many applications, $\vec{x}^T A\vec{x}$ is associated with the energy in the system. This is why we call it an energy criterion.

This criterion leads to a very nice, very easy result:

If A and B are symmetric positive definite, then so is $A+B$.

Indeed, $\vec{x}^T(A+B)\vec{x} = \vec{x}^T A\vec{x} + \vec{x}^T B\vec{x}$

Finally, consider a 2x2 symmetric positive definite matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Then for any $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $[x \;\; y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} > 0$

$$\Leftrightarrow ax^2 + 2bxy + cy^2 > 0$$

We see that the graph of a 2×2 symmetric positive definite matrix is an ellipsoidal bowl going up from $\vec{x} = (0,0)$.

4) Factorization criterion

Consider a matrix R with independent columns (R may be rectangular). $A = R^T R$ is square and symmetric. Let us show it is also positive definite:

For any $\vec{x} \neq \vec{0}$, $\vec{x}^T A \vec{x} = (R\vec{x})^T R\vec{x} = \|R\vec{x}\|^2 > 0$

$R\vec{x} = \vec{0}$ is not possible since the columns of R are independent.

Conversely, any symmetric positive definite matrix A can be written as $R^T R$ with independent columns for R. Indeed, consider the LDU decomposition for A, which is symmetric:

$$A = L D L^T$$

Now, if A is positive definite, the diagonal entries of D, which are the pivots of A, are all positive. We can therefore define $\sqrt{D}$, the diagonal matrix whose nonzero entries are the square roots of the pivot.

$$A = (L\sqrt{D})(L\sqrt{D})^T = R^T R, \text{ with } R = (L\sqrt{D})^T$$

Another option is to consider the diagonalization $A = Q \Lambda Q^T$, and choose $R = Q\sqrt{\Lambda}\, Q^T$.

Bottom line: A symmetric matrix A is positive definite if and only if $A = R^T R$ for a matrix R with independent columns.

Example: $\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{3} & 0 \\ \sqrt{3} & \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{2} \end{bmatrix} = R^T R$$

with $R = \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{2} \end{bmatrix}$

II] Symmetric positive semidefinite matrices

Some matrices have all their eigenvalues positive, but not strictly positive. In other words, $\lambda = 0$ is an eigenvalue for these matrices, which therefore are singular.

Another way to put it is that for these matrices, for all $\vec{x}$, $\vec{x}^T A \vec{x} \geq 0$, and there are nonzero $\vec{x}$ such that $\vec{x}^T A \vec{x} = 0$.

Such matrices are called positive semidefinite.

Example. $A = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$ has 2 and 0 as pivots, so it has a strictly positive eigenvalue, and also 0 as an eigenvalue. It is positive semidefinite

III Illustration: Axes of an ellipse, eigenvectors, and eigenvalues

Consider a symmetric positive definite matrix A with dimensions 2×2:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

$$\vec{x}^T A \vec{x} = ax^2 + 2bxy + cy^2 \quad \text{so} \quad \vec{x}^T A \vec{x} = 1$$
$$\Leftrightarrow ax^2 + 2bxy + cy^2 = 1$$

This is the equation of a tilted ellipse

Since A is symmetric, we may write $A = Q \Lambda Q^T$

And since A is positive definite, $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ with $\lambda_1 > 0, \lambda_2 > 0$

Then, $$\vec{x}^T A \vec{x} = \vec{x}^T Q \Lambda Q^T \vec{x}$$
$$= (Q^T \vec{x})^T \Lambda Q^T \vec{x}$$
$$= \vec{X}^T \Lambda \vec{X} = 1$$

with $\vec{X} = Q^T \vec{x} = \begin{bmatrix} X \\ Y \end{bmatrix}$

Now, $\vec{X}^T \Lambda \vec{X} = \lambda_1 X^2 + \lambda_2 Y^2 = 1$

So in the new coordinates $\vec{X} = Q^T \vec{x}$, the ellipse is upright, with axis half-lengths $\frac{1}{\sqrt{\lambda_1}}$ and $\frac{1}{\sqrt{\lambda_2}}$

The matrix which rotates the axes x and y so they align with the ellipse is Q^T.

The axes of the tilted ellipse point along the eigenvectors of A, which are (when normalized to have length 1) the columns of Q.

Example: Find the axes of the tilted ellipse

$4x^2 + 2xy + 4y^2 = 1$

The associated matrix is $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

Its eigenvalues are such that $(4-\lambda)^2 - 1 = 0$

$$\Leftrightarrow (3-\lambda)(5-\lambda) = 0$$

$\lambda = 3$ and $\lambda = 5$ are the eigenvalues.

Its eigenvectors are in the nullspace of $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$, so $\vec{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which give the axes of the ellipse.

Orthonormal vectors are then easily found:

$q_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\vec{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, so we may write:

$$\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Multiplying by $[x\ y]$ on the left and $\begin{bmatrix} x \\ y \end{bmatrix}$ on the right, we have

$$\vec{x}^T A \vec{x} = 3\left(\frac{x-y}{\sqrt{2}}\right)^2 + 5\left(\frac{x+y}{\sqrt{2}}\right)^2 = 1$$

We see that the ellipse is upright in the coordinate system such that $\frac{x+y}{\sqrt{2}} = X$ and $\frac{x-y}{\sqrt{2}} = Y$.