2.3 Symmetric Positive Definite Matrices

In this section we will restrict the matrices to symmetric matrices and then to symmetric positive definite matrices. Although this restriction may seem a little severe, there are a number of important applications, which include some classes of partial differential equations and some classes of least squares problems. The advantage of this restriction is that the number of operations to do Gaussian elimination can be cut in half.

Definition. A is a **symmetric** matrix if $A = A^T$.

Definition. **Cholesky** factorization of A is $A = \hat{L}\hat{L}^T$ where $\hat{L}$ is a low triangular matrix.

Proposition 4. If $A = A^T$, $a_{kk}^{(k)} \neq 0$ and $A = LU$, then $A = LDL^T$, where

$$D \quad \mathrm{diag}(a_{11}^{(1)}, a_{22}^{(2)},...,a_{nn}^{(n)}),\ k=1,2,...,n$$

Proof: $A = LU = LDD^{-1}U = LDM^T$ and here $M^T \quad D^{-1}U$ is an upper triangular matrix with ones on the diagonal. Note the following:

(1). $M^{-1}AM^{-T} = M^{-1}(LU)(D^{-1}U)^{-1}$

$$= M^{-1}LUU^{-1}D$$

$$= M^{-1}LD$$

$M^{-1}L$ is the product of low triangular matrices with ones on the diagonal, and so, $M^{-1}LD$ is a lower triangular matrix with diagonal part equal to D.

(2). $(M^{-1}AM^{-T})^T = (M^{-T})^T A^T (M^{-1})^T$

$= M^{-1}AM^{-T}$, because A is symmetric.

So, $M^{-1}LD$ is a symmetric matrix.

(3). By (1) and (2) $M^{-1}LD = D$ so that $M^{-1}L = I$ and $M = L$.

Note: If diagonal components of D are positive, then $D=D^{1/2}D^{1/2}$, and

$$A = LDL^T = (LD^{1/2})(D^{1/2}L^T) = (LD^{1/2})(LD^{1/2})^T) = \hat{L}\hat{L}^T .$$

The following additional restriction will ensure this is true.

Definition. A is **real positive definite** if and only if for all $x \neq 0$, $x^TAx > 0$

Examples.

1. $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

   $x^TAx = x_1^2 + (x_1-x_2)^2 + x_2^2 > 0$

   A similar nxn matrix with

   $$\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$$ is positive definite.

2. $A = A^T$ with $a_{ii} > 0$ and $a_{ii} > \sum_{j \neq i} |a_{ij}|$ is positive definite.

3. Finite element method for elliptic PDE (partial differential equation)

   $-\Delta U + c U = f$ where c is positive and U is zero on the boundary

Convert it into weak form by multiplying φ, where this function is zero on the boundary, and using Green's Theorem:

$$\int_\Omega \nabla U \, \nabla\varphi + c\, U\varphi \;=\; \int_\Omega f\varphi$$

For two space dimensions,

$$\int_\Omega U_x\, \varphi_x \;+\; U_y\, \varphi_y + c\, U\varphi = \int_\Omega f\varphi$$

Let $U \cong \sum_{j=1}^n u_j\, \varphi_j(x,y)$, $u_j \in \Re$, j=1,2,...,n and

$$\varphi \;\rightarrow\; \varphi_i(x,y), \;\; \varphi_i(x,y) \text{ is a given function.}$$

$$\int((\textstyle\sum_j u_j\, \varphi_j)_x\varphi_{ix} + (\sum_j u_j\, \varphi_j)_y\varphi_{iy} + (\sum_j u_j\, \varphi_j)\, \varphi_I) = \int f\varphi_i, \quad i=1,2,..., n$$

$$\textstyle\sum_j \int(\varphi_{jx}\varphi_{ix} + \varphi_{jy}\varphi_{iy} + c\, \varphi_j\varphi_i)\, u_j \;=\; \int f\varphi_i$$

Let $a_{ij} = \int(\varphi_{jx}\varphi_{ix} + \varphi_{jy}\varphi_{iy} + c\, \varphi_j\varphi_i)$ so that

$$u^T A u = \textstyle\sum_i\sum_j a_{ij}u_iu_j$$

$$= \int((\textstyle\sum_j u_j\, \varphi_j)_x^{\,2} + (\sum_j u_j\, \varphi_j)_y^{\,2} + c\, (\sum_j u_j\, \varphi_j)^2) > 0.$$

4. Normal equation from least squares.

Let A be mxn where m > n and consider A x = d, or

r(x) = d - Ax = 0. Since m > n, there may be not solution.

The **least square problem** is to find x such that

$$r(x)^T r(x) = \min_y \; r(y)^T r(y)$$

If A has full column rank (Ax = 0 implies x = 0), then the **normal equation** $A^TAx = A^Td$ is equivalent to finding the least squares solution. Here A^TA is SPD (symmetric positive definite) because:

$$\text{if } x \neq 0 \text{, then } Ax \neq 0,$$

$$x^T(A^TA)x = (Ax)^T(Ax) = \| Ax \|_2^2 > 0 .$$

Proposition 5. If A is positive definite, then

1. A^{-1} exists.

2. $a_{ii} > 0$.

3. Principle submatrices are also positive definite.

4. If S has full column rank, then S^TAS is positive definite.

5. If A is SPD, then eigenvalues of A (a complex number λ such that $Ax = \lambda x$, $x \neq 0$) are real and positive.

6. Let M be symmetric matrix and

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

M is positive definite if and only if A and $C - B^TA^{-1}B$ are positive definite.

Proofs:

1. We need to show $Ax = 0$ implies $x = 0$.

   Suppose $Ax = 0$ and $x \neq 0$. Because A is positive definite $x^TAx \neq 0$.

   But, this contradicts $Ax = 0$.

2. Let $x = e_i$, $0 < e_i^TAe_i = a_{ii}$

6. First, suppose that M is SPD. Use part three applied to A in M to show A is also SPD and nonsingular. The M can be factored as follows:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ (A^{-1}B)^T & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}.$$

Hence, by part four the diagonal matrix on the right side must be SPD

$$\begin{bmatrix} I & 0 \\ -(A^{-1}B)^T & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix}.$$

Then the second diagonal block must also be SPD.

Second, suppose both the first and second diagonal blocks in the above right side are SPD. In order to show M is SPD, just use the above block factorization of M and apply part four.

Proposition 6. If A is SPD, then A = $\hat{L}\hat{L}^T$, where $\hat{L}$ $LD^{1/2}$, $a_{kk}^{(k)} > 0$.

Proof: Since A is positive definite, $a_{11}^{(1)} = a_{11} > 0$. So, we may apply the row operations to the first column

$$(\prod E_{i1})\, A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

Next, apply the corresponding column operations to the first row, and use the symmetry of A to get

$$(\prod E_{i1})\, A (\prod E_{i1})^T = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

Thus, the lower diagonal (n-1)x(n-1) block must also be SPD, and one can repeat the above arguments on this smaller block to get $a_{22}^{(2)} > 0$. A more formal proof should use mathematical induction on n.