Fall-2023 5304 LecN6 Notes

Wan

November 8, 2023

Topics: PD; SPD; LDLT; Cholesky Factorization.

Math Tools: Inner Product, Eigenvalues and Eigencevectors, Properties of some special matrices

1 Positive Definite Matrices

Definition of PD

A real, square matrix is said to be PD if:
$\langle Au, u \rangle > 0$ for all $u \neq 0$ and $u \in R^n$

1.1 Properties

1, A is nonsingular.
This can be proof by contradiction.

2, The eigenvalues of A are real and positive.
This can be proved by the definition of eigenvalues.

2 Pred: SPD, Semi-Definite, Neg definite, and Indefinite Matrices

2.1 Pred by the quadratic form

Pred steps:
1, transform the matrix to the quadratic form, method as below:

①含有n个变量$x_1, x_2, \cdots x_n$的二次齐次多项式 $f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n + 2a_{23}x_2x_3 + \cdots + 2a_{n-1}x_{n-1}x_n$ 称为二次型.

②二次型对应的矩阵为:

下标和项的下标一致.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

例: $$A = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 9 & 6 \\ 6 & a \end{bmatrix} \end{array}$$

$$x^TAx = a_{11}x_1^2 + a_{22}x_2^2 + a_{12}x_1x_2 + a_{21}x_1x_2$$
$$= 9x_1^2 + 6x_1x_2 + 6x_1x_2 + ax_2^2$$
$$= 9x_1^2 + 12x_1x_2 + ax_2^2.$$

2, pred by this:

对于二次型函数，$f(x)=x^TAx$：

- $f(x)>0, x\neq 0, x\in\mathbb{R}$，则 f 为正定二次型，A 为正定矩阵
- $f(x)\geq 0, x\neq 0, x\in\mathbb{R}$，则 f 为半正定二次型，A 为半正定矩阵
- $f(x)<0, x\neq 0, x\in\mathbb{R}$，则 f 为负定二次型，A 为负定矩阵
- $f(x)\leq 0, x\neq 0, x\in\mathbb{R}$，则 f 为半负定二次型，A 为半负定矩阵
- 以上皆不是，就叫做不定

Example

Example

$$A = \begin{bmatrix} 9 & 6 \\ 6 & a \end{bmatrix}$$

$$x^TAx = 9x_1^2 + 12x_1x_2 + ax_2^2 = (3x_1+2x_2)^2 + (a-4)x_2^2$$

- A is positive definite for $a>4$

$$x^TAx > 0 \quad \text{for all nonzero } x$$

- A is positive semidefinite but not positive definite for $a=4$

$$x^TAx \geq 0 \quad \text{for all } x, \qquad x^TAx = 0 \quad \text{for } x=(2,-3)$$

- A is not positive semidefinite for $a<4$

$$x^TAx < 0 \quad \text{for } x=(2,-3)$$

2.2 Pred by leading principal minors

定理1：正定性检验

设 A 是一个 $n \times n$ 对称矩阵。则：

（a） A 是正定矩阵，当且仅当 A 的 n 个顺序主子式(严格)为正。

（b） A 是负定矩阵，当且仅当 A 的 n 个顺序主子式以如下方式交替出现：

> $|A_1| < 0$， $|A_2| > 0$， $|A_3| < 0$，……

即 k 阶顺序主子式的符号应该与 $(-1)^k$ 相同。

（c）如果 A 的某个 k 阶顺序主子式非零，但不符合上述两个符号模式中的任何一个，则 A 是不定的。当 A 对偶数 k 有一个负 k 阶的顺序主子式，或者当 A 对两个不同的奇数 k 和 ℓ 有一个负 k 阶的顺序主子式和一个正 ℓ 阶的顺序主子式时，就会发生这种情况。

对于给定对称矩阵 A，定理1的顺序主子式检验可能失败。一种原因是， A 的一些顺序主子式为零，而非零的其他顺序主子式则符合定理1中（a）或（b）的符号模式。当这种情况发生时，矩阵A是不定的，它可能是半定，也可能不是半定。在这种情况下，为了检查半定性，不再只检查 A 的 n 个顺序主子式，而是必须检查 A 的每个主子式的符号，使用下面的定理描述的检验。

定理2：半定性检验

设 A 是一个 $n \times n$ 对称矩阵。则：

- A 是正半定的，当且仅当 A 的每一个主子式都大于等于0
- A 是负半定的，当且仅当 A 的每一个奇阶的主子式小于等于0，每一个偶阶的主子式大于等于0

2.3 Conclusions

1, Semi-definite matrix may be nonsingular, not as positive definite one.

3 Gram matrix $X^T X$

Gram matrix is a special type of matrix to describe the inner product relationships between sets of vectors. Specifically, if there is a set of vectors $[v1, v2, ..., vn]$, then the elements G_{ij} of the Gram matrix G are the inner products of vectors v_i and v_j, that is, $Gij =< vi, vj >$. The Gram matrix is semi-positive definite, and it is positive definite when the set of vectors is linearly independent.

Gram matrix

recall the definition of *Gram matrix* of a matrix B (page 4.20):

$$A = B^T B$$

- every Gram matrix is positive semidefinite

$$x^T A x = x^T B^T B x = \|Bx\|^2 \geq 0 \quad \forall x$$

- a Gram matrix is positive definite if

$$x^T A x = x^T B^T B x = \|Bx\|^2 > 0 \quad \forall x \neq 0$$

  in other words, B has linearly independent columns

4 Symmetric positive definite matrices, SPD

4.1 Definition of SPD

A square matrix is SPD **iff** it is **symmetric** and all its eigenvalues λ are **positive**, that is $\lambda > 0$.
Note:
- Don't forget the 2nd part of the definition, sometimes useful to think in this way.
- This is an iff relationship.

Properties

1, Diagonal entries of SPD is positive.

Proof starting from the definition: Known that A is SPD, then for all nonzero vector u, we have $\langle Au, u \rangle > 0$. Then, utilize identity matrix to extract an element of A. (MV Product, dot product view). Good enough.

2, SPD iff Each A_k is SPD
will fill this part later.

3, a conclusion
will fill this part later.

4, If A is SPD, then for any nxk matrix X of rank k, the matrix X^TAX is SPD.
Memory:

$k < n$

$$\underbrace{\begin{bmatrix} \quad X^T \quad \end{bmatrix}}_{k \times n} \underbrace{\begin{bmatrix} A \end{bmatrix}}_{n \times n} \underbrace{\begin{bmatrix} X \end{bmatrix}}_{n \times k}$$

Proof (will use the notation of the inner product and def of SPD):

Application: Covariance Matrices in Statistics

5 The LDL^T Factorization from LU

Theorem

Theorem. *Let S be a positive-definite symmetric matrix. Then S has unique decompositions*

$$S = LDL^T \quad \text{and} \quad S = L_1 L_1^T$$

where:

- *L is lower-unitriangular,*
- *D is diagonal with positive diagonal entries, and*
- *L_1 is lower-triangular with positive diagonal entries.*

For the LU Factorization, all we need is all A_k matrices for k=1 to n-1 have to be nonsingular. As A is SPD, implying that all A_k matrices from k=1 to n are nonsingular. Thus LU fac exists and is unique.

Inverse of a Diagonal Matrix is easy to compute:

Inverse of a Diagonal Matrix

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \qquad D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}$$

Nice results because of symmetry

L is M

Proof that $A = LU = LDL^T$

Thus, a SPD matrix A could be written in this form: $A = LDL^T$ where L is a lower triangular matrix with 1s on the diagonal, and D is a diagonal matrix (of U).

6 The Cholesky Factorization GGT from LDLT

GGT, scope of application: SPD matrices (??? what about semi-definte matrix???).

Diagonal of D shoud be positive so that we can take a square root to D to reach the Cholesky.

Remark. Suppose that S has an LDL^T decomposition with

$$D = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}.$$

Then we define

$$\sqrt{D} = \begin{pmatrix} \sqrt{d_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{d_n} \end{pmatrix},$$

so that $(\sqrt{D})^2 = D$, and we set $L_1 = L\sqrt{D}$. Then

$$L_1 L_1^T = L(\sqrt{D})(\sqrt{D})^T L^T = LDL^T = S,$$

so $L_1 L_1^T$ is the Cholesky decomposition of S.

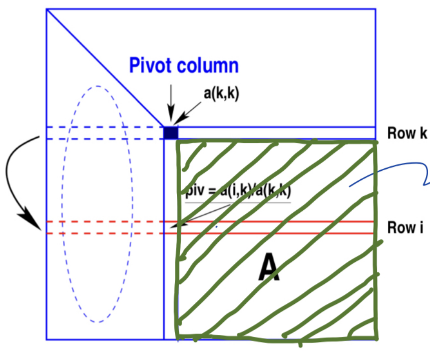
Proof: Diagonal of D shoud is positive

What can we say about G? G is lower triangular
The Cholesky: Any matrix that is SPD could be written as $A = GGT$ where G is a lower triangular matrix, and positive entries on diagonal.

7 Algos of LDLT and GGT

row-oriented LDLT

Idea: Adapted from Gaussian but just work on upper part of the matrix because of symmetry because the working matrix A(k+1:n, k+1:n) in standard LU remains symmetric. See graph below:



remains symmetric.

WHY?

Proof: A(k+1:n, k+1:n) is symmetric (induction argument).

What is the cost(FLOPs) of this algo?
Cost of standard LU:
Suppose k's index starting from 0, then in standard LU decomposition, when clearing the 1st col, each row replacement invoves n-1 multiplications (scale piv row) and n-1 substraction (update curr row), 2(n-1) flops in total. Hence, when k increase, problem size decrese, it takes $2(n-2)^2$ flops to clear the 2nd col etc.. Thus, we have:

$$(n-1)\cdot 2(n-1) + (n-2)\cdot 2(n-2) + \dots + 1$$
$$= 2\left((n-1)^2 + (n-2)^2 + \dots + 1\right) = 2\cdot \frac{n(n-1)(2n-1)}{6} \approx \frac{2}{3}n^3.$$

平方和求和公式.

平方和与立方和

平方和与立方和的求和公式如下：

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \qquad \text{(A. 3)}$$

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4} \qquad \text{(A. 4)}$$

Cost of LDLT (just half of LU): $\frac{1}{2} * \frac{2}{3} * n^3 = \frac{1}{3} * n^3$

Rank 1 update: You're updating the matrix by something like uv^T where u and v are the same.

2, row GGT, outer product form

3, column-oriented LDLT