

# ST 661 Note

Frances Lin

1/6/2022

## Outline

### Woodbury Matrix Identity

If

$$A = B + CC^T,$$

then

$$A^{-1} = (B + CC^T)^{-1} = B^{-1} - \frac{B^{-1}CC^TB^{-1}}{1 + C^TB^{-1}C},$$

where  $B^{-1}$  is a familiar matrix so no new inversion is needed.

### (Generalized) Woodbury Matrix Identity

More generally,

$$(A + PBQ)^{-1} = A^{-1} - A^{-1}PB(B + BQA^{-1}PB)^{-1}BQA^{-1}.$$

e.g.

$$y = X\beta + \varepsilon,$$

where

$$\text{cov}(\varepsilon) = \sigma_1^2 I + \sigma_2^2 K = \Sigma,$$

where  $\sigma_1^2 I$  is the identify term and  $\sigma_2^2 K$  is the correlation term.  $\sigma_1^2 I$  captures measurement error, for example, and  $\sigma_2^2 K$  captures dependence.  $K$  is a square matrix and may be low rank.

For this mixed effect model, e.g.

$$K = \begin{pmatrix} 1 & \cdots & 1 & & & & \\ 1 & \cdots & 1 & & 0 & & 0 \cdots 0 \\ 1 & \cdots & 1 & & & & \\ & 0 & & 1 & \cdots & 1 & \\ & & & 1 & \cdots & 1 & \\ & & & 1 & \cdots & 1 & \\ & 0 & & & & & \ddots \\ & \vdots & & & & & \\ & 0 & & & & & \end{pmatrix}$$

is a block matrix with the first 2 blocks shown.  $\text{rank}(K) =$  the number of blocks in the experiment.

Our goal is to invert  $\Sigma$  and invert it quickly. Suppose

$$\text{rank}(K) = r \ll n,$$

then we can write

$$K = ZZ^T,$$

where  $K$  is a square matrix,  $Z$  is column matrix ( $n \times r$ ) and  $Z^T$  is a row matrix ( $r \times n$ ).

Then,

$$\Sigma^{-1} = (\sigma_1^2 I + \sigma_2^2 K)^{-1} = (\sigma_1^2 I + \sigma_2^2 ZZ^T)^{-1} = \frac{1}{\sigma_1^2} I - \frac{\sigma_2^4}{\sigma_1^4} Z(\sigma_2^2 I + \frac{\sigma_2^2}{\sigma_1^2} Z^T Z)^{-1} Z^T,$$

where  $ZZ^T$  is a low rank modification and  $\sigma_2^2 I + \frac{\sigma_2^2}{\sigma_1^2} Z^T Z$  is  $r \times r$  (smaller dimension).

We will discuss inverse of partitioned matrix later.

## Positive Definite Matrix

### Quadratic Equation

e.g.  $ax^2 + bx + c$

### Quadratic Form

If we have more than 1  $x$ , we call it quadratic form.

e.g.

$$3y_1^2 + y_2^2 + 2y_3^2 + 4y_1y_2 + 5y_1y_3 - 6y_2y_3 \quad (*)$$

This is a homogeneous polynomial of degree 2. Right now, nothing is random here.

If we let

$$y = (y_1, y_2, y_3)^T$$

be a column vector, then

$$(*) = y^T A y,$$

where, for example,  $y^T$  is  $1 \times 3$ ,  $A$  is  $3 \times 3$ , and  $y$  is  $3 \times 1$ . One possibility of  $A$  is that

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & -6 \\ 0 & 0 & 2 \end{bmatrix}.$$

Another possibility of  $A$  is that

$$B = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & -6 \\ 4 & 0 & 2 \end{bmatrix}.$$

Notice that  $y^T A y = y^T B y$  but  $A \neq B$ . Notice also that diagonal elements are the same. However, off diagonal elements are different. Indeed, in  $A$ , 5 from  $(*)$  and  $A$  is split between 1 and 4. Matrix  $A^*$  is not unique.

We want to constrain  $A$  s.t.  $A^* = C$  is symmetric, then this matrix is unique.

$$A^* = C = \begin{bmatrix} 3 & 2 & 2.5 \\ 2 & 1 & -3 \\ 2.5 & -3 & 2 \end{bmatrix}.$$

This is obtained by dividing the coefficients of (\*) for  $4y_1y_2 + 5y_1y_3 - 6y_2y_3$  by 2 (i.e.  $4/2 = 2$ ,  $5/2 = 2.5$ , etc). Indeed,

$$C = \frac{AA^T}{2}$$

makes it a symmetric matrix.

In general,  $y^T Cy = y^T Ay$ . It suffices to study  $y^T Ay$  for symmetric  $A$ .

## Positive Definite Matrix

Positive definite matrix is considered a special case of symmetric matrix.

If a symmetric matrix  $A$  is s.t.  $y^T Ay > 0$  for all possible  $y$ s except  $y = 0$ , then the quadratic form is said to be positive definite and  $A$  is said to be a positive definite (pd) matrix.

## Positive Semi-Definite Matrix

If a symmetric matrix  $A$  is s.t.  $y^T Ay \geq 0$  for all possible  $y$ s except  $y = 0$ , then the quadratic form is said to be positive definite and  $A$  is said to be a positive semi-definite (psd) matrix.

e.g.

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$$

is symmetric. There are many tool to check but now we stick to the definition.

$$y^T Ay = 2y_1^2 - 2y_1y_2 + 3y_2^2 = 2\left(y_1 - \frac{1}{2}y_2\right)^2 + \frac{5}{2}y_2^2 \geq 0$$

To make  $y^T Ay = 0$ , then

$$y_1 - \frac{1}{2}y_2 = 0 \Leftrightarrow y_1 = y_2 = 0$$

and

$$\frac{5}{2}y_2^2 = 0 \Leftrightarrow y_2 = 0.$$

$y_1$  and  $y_2$  is trivial. Therefore,  $A$  is pd.

e.g.

$$B = \begin{bmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix}$$

is psd.

$$y^T Ay = (2y_1 - y_2)^2 + (3y_1 - y_3)^2 + (3y_2 - 2y_3)^2 \geq 0$$

To make  $y^T Ay = 0$ , then

$$2y_1 = y_2, 3y_1 = y_3, 3y_2 = 2y_3$$

$$\Leftrightarrow y = (1, 2, 3)^T$$

$\exists$  nontrivial  $y$  to make this zero, so  $B$  is not pd.  $B$  is psd, however.

## Thm

- i) If  $A$  is pd (positive definite), then all diagonal elements  $a_{ii}$  are positive.
- ii) If  $A$  is psd (positive semi-definite), then all elements  $a_{ii}$  are nonnegative.

Proof.

$$y^T A y > 0, y \neq 0$$

$$y = [0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0]^T$$

where 1 occurs at the  $i^{th}$  spot, then

$$y^T A y = a_{ii} > 0 \text{ for pd}$$

The reason is that  $y^T$  extracts the  $i^{th}$  row and  $y$  extracts the  $i^{th}$  column. Having both  $y^T$  and  $y$  extract the  $ii$  entry.

Similar, we can show that

$$y^T A y = a_{ii} \geq 0 \text{ for psd.} \quad \square$$

Ques. What if we want to work with a partitioned matrix?

For example,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

It does not matter the size of  $A_{11}$  and  $A_{22}$  as long as they are both square matrices.

## Thm

If  $A$  is pd of the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where  $A_{11}$  and  $A_{22}$  are both square, then  $A_{11}$  and  $A_{22}$  are pd.

Tentative proof. Suppose  $BAB^T = A^{11}$  and

$$B = [I \quad 0] = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

retains the first few rows. On the other hand,  $B^T$  retains the first few columns. Together,

$$BAB^T = A_{11}.$$

Next, we need to argue that  $BAB^T$  is pd when  $A$  is pd. We will need a Lemma.

## Thm

Let  $P$  be a nonsingular (square) matrix,

- i) If  $A$  is pd, then  $P^T A P$  is pd.
- ii) If  $A$  is psd, then  $P^T A P$  is psd.

Proof. We need to show that

- i)  $y^T (P^T A P) y > 0$  whenever  $y^T \neq 0$

Since

$$y^T (P^T A P) y = (Py)^T A (Py),$$

and  $P$  is nonsingular, so when  $y^T \neq 0$ ,  $Py \neq 0$ . Next, given that  $A$  is pd,

$$(Py)^T A (Py) > 0. \quad \square$$

ii) Exercise.

## Lemma - Generalized

Suppose  $B$  is  $k \times p$ , then there are two cases:

- i)  $k \leq p$  so a row matrix
- ii)  $k > p$  so a column matrix

In case i),  $BAB^T$  results in a smaller square matrix and is *possible* pd.

In case ii),  $BAB^T$  results in a larger square matrix and is no longer pd.

$k \times p$  times  $p \times p$  times  $p \times k = k \times k$

Proof. Suppose, by contradiction, that

$A$  is singular then  $\exists y$  s.t.  $Ay = 0$  for  $y \neq 0$ . Therefore,  $y^T Ay = 0$ . So,  $A$  is not pd, which contradicts.

## Corollary (Main Thm)

Let  $A$  be a  $p \times p$  pd matrix and  $B$  be a  $k \times p$  matrix,

- i) If  $\text{rank}(B) = k \leq p$  (full rank), then  $BAB^T$  is pd.

Refer to case i)

- ii) If  $k > p$  or  $\text{rank}(B) \leq \min(k, p)$  (not full rank), then  $BAB^T$  is psd.

Refer to case ii)

Recall from 01042022's note that  $A$  is said to be a full rank matrix if

$$\text{rank}(A) = \min(n, p).$$

End. Verify  $\Sigma^{-1}$ . Review Corollary and the proof above Corollary.