

# Thesis Contents

*Zhiwei Zhou*

August 30, 2022

Abstract

Acknowledgments

Contents

1 List of Symbols

2 List of Figures

3 List of Tables

4 Introduction

4.1 Basic Definition

4.2 Vector and Matrix Norm

4.3 Floating Point System

4.4 Floating Point Arithmetic

4.5 Matrix decompositions

4.6 Forward an Backward Errors

5 Background(Literature review)

6 Derivation

7 Algorithms

7.1 Improvement(expmodified)

8 Application

9 Numerical Expriment

10 Conclusion

11 Futrue Work

References

Appendix

[?]

## 12 Semiseparable

Define matrices  $E_u(p)$  and  $E_l(q)$  as following,

$$E_u(p) = \begin{cases} 1, & \text{if } i \leq j - p, \\ 0, & \text{otherwise.} \end{cases} \quad \text{and} \quad E_l(q) = \begin{cases} 1, & \text{if } i \geq j - q, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

i.e.

$$\begin{aligned} E_u(-1) &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 1 & \ddots & \\ & 0 & \ddots & 1 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \quad E_u(0) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 0 & \ddots & \\ & \vdots & \ddots & 1 \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \quad E_u(1) = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 \end{bmatrix}; \\ E_l(-1) &= \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ & 1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1 & 1 \end{bmatrix}, \quad E_l(0) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ & 1 & \ddots & \\ \vdots & & \ddots & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix}, \quad E_l(1) = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \\ 1 & \cdots & 1 & 0 \end{bmatrix}. \end{aligned} \quad (2)$$

where  $|p| \leq n - 1$  and  $|q| \leq n - 1$ .

Define the operations triu and tril,

$$\text{triu}(A, p) = A \circ E_u(p) \quad (4)$$

$$\text{tril}(A, q) = A \circ E_l(q) \quad (5)$$

i.e.  $\text{triu}(A, p)$  retains the entries of  $A$  above the  $p$ -th super-diagonal (inclusive) and makes the entries of  $A$  below the  $p$ -th super-diagonal (exclusive) 0.

define symmetric  $\{1,1\}$ -semiseparable matrix,  $S$

$$S = \begin{cases} x_i y_j & \text{if } i \leq j; \\ x_j y_i & \text{if } i \geq j. \end{cases} \quad (6)$$

$$= \text{triu}(xy^T) + \text{tril}(yx^T, -1) \quad (7)$$

$$= \text{triu}(xy^T, 1) + \text{tril}(yx^T) \quad (8)$$

$$= \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_1 y_2 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 y_n & x_2 y_n & \cdots & x_n y_n \end{bmatrix} \quad (9)$$

## 13 Matrix (basic def)

### Matrix-Matrix Multiplication

$$AB = C \tag{10}$$

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \tag{11}$$

### Matrix-Vector Multiplication

$$Ax = b \tag{12}$$

$$b_i = \sum_{k=1}^n a_{ik}x_k \tag{13}$$

## 14 Algorithm (Appendix)

### Backward Substitution

$$Ux = b \quad (14)$$

### Triangular Matrix

$$\begin{bmatrix} d_1 & u_1^1 & u_1^2 & \cdots & \cdots & u_1^{n-1} \\ & d_2 & u_2^1 & u_2^2 & \cdots & u_2^{n-2} \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & d_{n-2} & u_{n-2}^1 & u_{n-2}^2 \\ & & & & d_{n-1} & u_{n-1}^1 \\ & & & & & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-2} \\ b_{n-1} \\ b_n \end{bmatrix} \quad (15)$$

$$x_n = b_n / u_{nn} \quad (16)$$

$$x_i = \left( b_i - \sum_{j=i+1}^n u_{ij} x_j \right) / u_{ii} \quad (17)$$

$$d_n x_n = b_n \quad \Rightarrow \quad x_n = b_n / d_n \quad (18)$$

$$d_{n-1} x_{n-1} + u_{n-1}^1 x_n = b_{n-1} \quad \Rightarrow \quad x_{n-1} = (b_{n-1} - u_{n-1}^1 x_n) / d_{n-1} \quad (19)$$

$$d_{n-2} x_{n-2} + u_{n-2}^1 x_{n-1} + u_{n-2}^2 x_n = b_{n-2} \quad \Rightarrow \quad x_{n-2} = (b_{n-2} - u_{n-2}^1 x_{n-1} - u_{n-2}^2 x_n) / d_{n-2} \quad (20)$$

$$\vdots \quad (21)$$

$$b_{n-k} = d_{n-k} x_{n-k} + u_{n-k}^1 x_{n-k+1} + u_{n-k}^2 x_{n-k+2} + \cdots + u_{n-k}^k x_{n-k+k} \quad (22)$$

$$\Rightarrow x_{n-k} = \left( b_{n-k} - u_{n-k}^1 x_{n-k+1} - u_{n-k}^2 x_{n-k+2} - \cdots - u_{n-k}^k x_{n-k+k} \right) / d_{n-k} \quad (23)$$

$$= \left( b_{n-k} - \sum_{j=1}^k u_{n-k}^j x_{n-k+j} \right) / d_{n-k} \quad (24)$$

$$(25)$$

$$x_i = \left( b_i - \sum_{j=1}^{n-i} u_i^j x_{i+j} \right) / d_i \quad (26)$$

$$x_i = \left( b_i - \sum_{j=i+1}^n u_{ij} x_j \right) / u_{ii} \quad (27)$$

## Upper Banded Matrix

$$\begin{bmatrix} d_1 & b_1^1 & b_1^2 & \cdots & b_1^u & & \\ & d_2 & b_2^1 & b_2^2 & \cdots & b_2^u & \\ & & \ddots & \ddots & \ddots & & \ddots \\ & & & \ddots & \ddots & \ddots & b_{n-u}^u \\ & & & & \ddots & \ddots & \vdots \\ & & & & & d_{n-2} & b_{n-2}^1 & b_{n-2}^2 \\ & & & & & d_{n-1} & b_{n-1}^1 \\ & & & & & & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ \vdots \\ b_{n-2} \\ b_{n-1} \\ b_n \end{bmatrix} \quad (28)$$

$$x_n = b_n / u_{nn} \quad (29)$$

$$x_i = \left( b_i - \sum_{j=i+1}^n u_{ij} x_j \right) / u_{ii} \quad (30)$$

$$d_n x_n = b_n \quad \Rightarrow \quad x_n = b_n / d_n \quad (31)$$

$$d_{n-1} x_{n-1} + u_{n-1}^1 x_n = b_{n-1} \quad \Rightarrow \quad x_{n-1} = (b_{n-1} - u_{n-1}^1 x_n) / d_{n-1} \quad (32)$$

$$d_{n-2} x_{n-2} + u_{n-2}^1 x_{n-1} + u_{n-2}^2 x_n = b_{n-2} \quad \Rightarrow \quad x_{n-2} = (b_{n-2} - u_{n-2}^1 x_{n-1} - u_{n-2}^2 x_n) / d_{n-2} \quad (33)$$

$$\vdots \quad (34)$$

$$b_{n-k} = d_{n-k} x_{n-k} + u_{n-k}^1 x_{n-k+1} + u_{n-k}^2 x_{n-k+2} + \cdots + u_{n-k}^k x_{n-k+k} \quad (35)$$

$$\Rightarrow x_{n-k} = \left( b_{n-k} - u_{n-k}^1 x_{n-k+1} - u_{n-k}^2 x_{n-k+2} - \cdots - u_{n-k}^k x_{n-k+k} \right) / d_{n-k} \quad (36)$$

$$= \left( b_{n-k} - \sum_{j=1}^k u_{n-k}^j x_{n-k+j} \right) / d_{n-k} \quad (37)$$

$$(38)$$

$$x_i = \left( b_i - \sum_{j=1}^{n-i} u_i^j x_{i+j} \right) / d_i \quad (39)$$

$$x_i = \left( b_i - \sum_{j=i+1}^n u_{ij} x_j \right) / u_{ii} \quad (40)$$

## 15 Apply on symmetric positive definite tridiagonal matrices

$$\longleftarrow \text{cholesky } \mathcal{O}\left(\frac{1}{3}n^3\right)$$

$$\begin{aligned} T &= U^T U \\ U^{-1} &= \text{triu}(xy^T) \\ (U^T)^{-1} &= (U^{-1})^T \\ &= (\text{triu}(xy^T))^T \\ &= \text{tril}(yx^T) \\ T^{-1} &= (U^T U)^{-1} \\ &= (U^{-1})(U^T)^{-1} \\ &= \text{triu}(xy^T) \text{tril}(yx^T) \end{aligned}$$

$$\begin{aligned} (T^{-1})_{i,j} &= \sum_{k=1}^n [\text{triu}(xy^T)]_{i,k} [\text{tril}(yx^T)]_{k,j} \\ &= \sum_{k=1}^n \left( \begin{cases} x_i y_k, & \text{if } i \leq j; \\ 0, & \text{if } i > j. \end{cases} \right) \left( \begin{cases} 0, & \text{if } k < j; \\ y_k x_j, & \text{if } k \geq j. \end{cases} \right) \\ &= \sum_{k=1}^n (x_i y_k \mathbb{1}_{\mathbb{Z} \cap [i,n]}(k)) (y_k x_j \mathbb{1}_{\mathbb{Z} \cap [j,n]}(k)) \\ &= x_i x_j \sum_{k=1}^n y_k^2 (\mathbb{1}_{\mathbb{Z} \cap [i,n]}(k) \mathbb{1}_{\mathbb{Z} \cap [j,n]}(k)) \\ &= x_i x_j \sum_{k=1}^n y_k^2 \mathbb{1}_{\mathbb{Z} \cap [\max(i,j),n]}(k) \end{aligned}$$

$$(T^{-1})_{i,j} = \begin{cases} x_i x_j \sum_{k=j}^n y_k^2, & \text{if } i \leq j; \\ x_j x_i \sum_{k=i}^n y_k^2, & \text{if } i \geq j. \end{cases} \quad (41)$$

for  $i = j$ ,

$$(T^{-1})_{i,i} = x_i^2 \sum_{k=i}^n y_k^2 = x_j^2 \sum_{k=j}^n y_k^2 = (T^{-1})_{j,j}$$

following the definition of  $\{1,1\}$ -Semiseparable matrix

$$S = \begin{cases} x_i y_j & \text{if } i \leq j; \\ x_j y_i & \text{if } i \geq j. \end{cases} \quad (42)$$

$$= \text{triu}(xy^T) + \text{tril}(yx^T, -1) \quad (43)$$

$$(44)$$

$$(T^{-1})_{i,j} = \begin{cases} x_i \left( x_j \sum_{k=j}^n y_k^2 \right), & \text{if } i \leq j; \\ x_j \left( x_i \sum_{k=i}^n y_k^2 \right), & \text{if } i \geq j. \end{cases} \quad (45)$$

$$= \begin{cases} \tilde{x}_i \tilde{y}_j & \text{if } i \leq j; \\ \tilde{x}_j \tilde{y}_i & \text{if } i \geq j. \end{cases} \quad (46)$$

$$= \text{triu}(\tilde{x} \tilde{y}^T) + \text{tril}(\tilde{y} \tilde{x}^T, -1) \quad (47)$$

$$(48)$$

where  $\tilde{x}_i = x_i$  and  $\tilde{y}_i = x_i \sum_{k=i}^n y_k^2$ .

$$\tilde{x}_i = x_i \quad (49)$$

$$\Rightarrow \tilde{x} = x \quad (50)$$

$$\tilde{y}_i = x_i \sum_{k=i}^n y_k^2 \quad (51)$$

$$\Rightarrow \tilde{y} = \begin{bmatrix} x_1 \left( \sum_{k=1}^n y_k^2 \right) \\ x_2 \left( \sum_{k=2}^n y_k^2 \right) \\ \vdots \\ x_n \left( \sum_{k=n}^n y_k^2 \right) \end{bmatrix} \quad (52)$$

$$= x \circ \begin{bmatrix} y^T y \\ y_{[2:n]}^T y_{[2:n]} \\ \vdots \\ y_n^2 \end{bmatrix} \quad (53)$$

$$\mathbb{1}\mathbb{1} \leq$$



## 16 28.06.2022 Meeting draft

### 16.1 Tridiagonal matrix (Upper)

$$n = 2k$$

$$\begin{aligned}
 U &= \begin{bmatrix} d_1 & b_1^1 & b_1^2 & & & & \\ & d_2 & b_2^1 & b_2^2 & & & \\ & & d_3 & b_3^1 & b_3^2 & & \\ & & & d_4 & b_4^1 & b_4^2 & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots \\ & & & & & & d_{n-3} & b_{n-3}^1 & b_{n-3}^2 \\ & & & & & & & d_{n-2} & b_{n-2}^1 & b_{n-2}^2 \\ & & & & & & & & d_{n-1} & b_{n-1}^1 & b_{n-1}^2 \\ & & & & & & & & & d_n & \\ & & & & & & & & & & d_n \end{bmatrix} \\
 &= \begin{bmatrix} \begin{bmatrix} d_1 & b_1^1 \\ & d_2 \end{bmatrix} & \begin{bmatrix} b_1^2 \\ b_2^1 & b_2^2 \\ d_3 & b_3^1 \end{bmatrix} & & & & \\ & \begin{bmatrix} b_3^2 \\ b_4^1 & b_4^2 \end{bmatrix} & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \begin{bmatrix} d_{2k-3} & b_{2k-3}^1 \\ & d_{2k-2} \end{bmatrix} & \begin{bmatrix} b_{2k-3}^2 \\ b_{2k-2}^1 & b_{2k-2}^2 \\ d_{2k-1} & b_{2k-1}^1 \end{bmatrix} & \\ & & & & \begin{bmatrix} b_{2k-1}^2 \\ d_{2k} \end{bmatrix} \end{bmatrix} \\
 &= \begin{bmatrix} D_1 & B_1 & & & \\ & D_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & D_{k-1} & B_{k-1} \\ & & & & D_k \end{bmatrix} \in \mathbb{R}^{2k \times 2k}
 \end{aligned}$$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{k-1} \\ X_k \end{bmatrix} \in \mathbb{R}^{2k \times 2} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{k-1} \\ Y_k \end{bmatrix} \in \mathbb{R}^{2k \times 2} \quad Y^T = [Y_1^T \ Y_2^T \ \dots \ Y_{k-1}^T \ Y_k^T] \in \mathbb{R}^{2 \times 2k}$$

$$\begin{aligned}
XY^T &= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{k-1} \\ X_k \end{bmatrix} [Y_1^T \ Y_2^T \ \cdots \ Y_{k-1}^T \ Y_k^T] \\
&= \begin{bmatrix} X_1 Y_1^T & X_1 Y_2^T & \cdots & X_1 Y_{k-1}^T & x_1 Y_k^T \\ X_2 Y_1^T & X_2 Y_2^T & \cdots & X_2 Y_{k-1}^T & x_2 Y_k^T \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ X_{k-1} Y_1^T & X_{k-1} Y_2^T & \cdots & X_{k-1} Y_{k-1}^T & X_{k-1} Y_k^T \\ X_k Y_1^T & X_k Y_2^T & \cdots & X_k Y_{k-1}^T & X_k Y_k^T \end{bmatrix} \\
&= X [Y_1^T \ Y_2^T \ \cdots \ Y_{k-1}^T \ Y_k^T] \\
&= [XY_1^T \ XY_2^T \ \cdots \ XY_{k-1}^T \ XY_k^T]
\end{aligned}$$

$$\begin{aligned}
\text{triu}(X_k) &:= X_k \\
\text{triu}(Y_k^T) &:= Y_k^T \\
\text{triu}(X_k Y_k^T) &= X_k Y_k^T
\end{aligned}$$

$$\begin{aligned}
\text{triu}(XY^T) &= \begin{bmatrix} \text{triu}(X_1 Y_1^T) & X_1 Y_2^T & \cdots & X_1 Y_{k-1}^T & X_1 Y_k^T \\ 0_{2 \times 2} & \text{triu}(X_2 Y_2^T) & \cdots & X_2 Y_{k-1}^T & X_2 Y_k^T \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{2 \times 2} & 0_{2 \times 2} & \cdots & \text{triu}(X_{k-1} Y_{k-1}^T) & X_{k-1} Y_k^T \\ 0_{2 \times 2} & 0_{2 \times 2} & \cdots & 0_{2 \times 2} & X_k Y_k^T \end{bmatrix} \\
&= \begin{bmatrix} \text{triu}(X_1 Y_1^T) \\ 0_{2(k-1) \times 2} \end{bmatrix} \begin{bmatrix} X_1 Y_2^T \\ \text{triu}(X_2 Y_2^T) \\ 0_{2(k-2) \times 2} \end{bmatrix} \cdots \begin{bmatrix} X_{1:k-2} Y_{k-1}^T \\ \text{triu}(X_{k-1} Y_{k-1}^T) \\ 0_{2 \times 2} \end{bmatrix} XY_k^T
\end{aligned}$$

$$\begin{aligned}
U^{-1} &= \text{triu}(XY^T) \\
I &= U \text{triu}(XY^T)
\end{aligned}$$

$$E_i = [0 \ \cdots \ 0 \ I_{2 \times 2} \ 0 \ \cdots \ 0]^T$$

$k$ -th block column,

$$\begin{aligned}
E_k &= U \text{triu}(XY^T) E_k \\
&= U X Y_k^T
\end{aligned}$$

$$\begin{aligned}
X &:= U^{-1} E_k \\
E_k &= U X Y_k^T \\
&= U U^{-1} E_k Y_k^T \\
&= E_k Y_k^T \\
Y_k^T &:= I_{2 \times 2}
\end{aligned}$$

$(k-1)$ -th block column,

$$\begin{aligned} E_{k-1} &= U \operatorname{triu}(XY^T)E_{k-1} \\ &= U \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \operatorname{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{2 \times 2} \end{bmatrix} \end{aligned}$$

$(k-1)$ -th block row of  $(k-1)$ -th block column,

$$E_i^T E_j = \begin{cases} I_{2 \times 2}, & \text{if } i = j ; \\ 0_{2 \times 2}, & \text{otherwise.} \end{cases}$$

$$\begin{aligned} E_{k-1}^T E_{k-1} &= I_{2 \times 2} = E_{k-1}^T U \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \operatorname{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{2 \times 2} \end{bmatrix} \\ &= \begin{bmatrix} 0_{2 \times 2(k-2)} & D_{k-1} & B_{k-1} \end{bmatrix} \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \operatorname{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{2 \times 2} \end{bmatrix} \\ &= D_{k-1} \operatorname{triu}(X_{k-1}Y_{k-1}^T) \\ Y_{k-1}^T &:= (D_{k-1}X_{k-1})^{-1} \end{aligned}$$

$$\begin{aligned} D_{k-1} &= \operatorname{triu}(D_{k-1}) \\ D_{k-1}^{-1} &= \operatorname{triu}(D_{k-1}^{-1}) \\ Y_{k-1}^T &:= (D_{k-1}X_{k-1})^{-1} \\ D_{k-1} \operatorname{triu}(X_{k-1}Y_{k-1}^T) &= D_{k-1} \operatorname{triu}(X_{k-1}(D_{k-1}X_{k-1})^{-1}) \\ &= D_{k-1} \operatorname{triu}(X_{k-1}X_{k-1}^{-1}D_{k-1}^{-1}) \\ &= D_{k-1} \operatorname{triu}(D_{k-1}^{-1}) \\ &= D_{k-1}D_{k-1}^{-1} \\ &= I_{2 \times 2} \end{aligned}$$

$$\begin{aligned} X &:= U^{-1}E_k \\ E_k &= UX \\ E_k^T E_k &= I_{2 \times 2} = E_k^T UX \\ &= \begin{bmatrix} 0_{2 \times 2(k-1)} & D_k \end{bmatrix} X \\ &= D_k X_k \\ Y_k^T &:= I_{2 \times 2} \\ &:= (D_k X_k)^{-1} \end{aligned}$$

$$\begin{aligned} X &:= U^{-1}E_k \\ Y_i^T &:= (D_i X_i)^{-1} \end{aligned}$$

$$n = 2k + 1$$

$$U = \begin{bmatrix} [d_1] & \begin{bmatrix} b_1^1 & b_1^2 \\ d_2 & b_2^1 \\ & d_3 \end{bmatrix} & \begin{bmatrix} b_2^2 & \\ b_3^1 & b_3^2 \end{bmatrix} & & \\ & & \ddots & \ddots & \\ & & & \begin{bmatrix} d_{2k-2} & b_{2k-2}^1 \\ & d_{2k-1} \end{bmatrix} & \begin{bmatrix} b_{2k-2}^2 & \\ b_{2k-1}^1 & b_{2k-1}^2 \\ d_{2k} & b_{2k}^1 \\ & d_{2k+1} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \tilde{D}_1 & \tilde{B}_1 & & & \\ & D_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & D_k & B_k \\ & & & & D_{k+1} \end{bmatrix} \in \mathbb{R}^{(2k+1) \times (2k+1)}$$

$$X := U^{-1} E_{k+1}$$

$$Y_i^T := (D_i X_i)^{-1} \text{ for } i = 2, \dots, k+1$$

$$Y_1^T := (D_1 X_1)^+ = (D_1 X_1)^T (D_1 X_1 (D_1 X_1)^T)^{-1}$$

## 16.2 Banded matrix (Upper)

$$n = ku + r, \quad 0 \leq r < u$$

$$U = \begin{bmatrix} d_1 & b_1^1 & b_1^2 & \cdots & b_1^u & & \\ & d_2 & b_2^1 & b_2^2 & \cdots & b_2^u & \\ & & \ddots & \ddots & \ddots & \cdots & \ddots \\ & & & \ddots & \ddots & \ddots & b_{n-u}^u \\ & & & & \ddots & \ddots & \vdots \\ & & & & & \ddots & b_{n-2}^2 \\ & & & & & & \ddots & b_{n-1}^1 \\ & & & & & & & d_n \end{bmatrix}$$

$$r = 0, \quad n = ku,$$

$$U = \begin{bmatrix} D_1 & B_1 & & & \\ & D_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & D_{k-1} & B_{k-1} \\ & & & & D_k \end{bmatrix} \in \mathbb{R}^{ku \times ku},$$

$$D_i \in \mathbb{R}^{u \times u}, \quad B_i \in \mathbb{R}^{u \times u} \text{ for } i = 1, 2, \dots, k.$$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix} \in \mathbb{R}^{ku \times u} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{bmatrix} \in \mathbb{R}^{ku \times u} \quad Y^T = [Y_1^T \quad Y_2^T \quad \cdots \quad Y_k^T] \in \mathbb{R}^{u \times ku}$$

$$X_i, Y_i^T \in \mathbb{R}^{u \times u}, \text{ for } i = 1, 2, \dots, k$$

$$X := U^{-1} E_k$$

$$Y_i^T := (D_i X_i)^{-1}$$

$$r \neq 0, \quad n = ku + r,$$

$$U = \begin{bmatrix} \tilde{D}_1 & \tilde{B}_1 & & & \\ & D_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & D_k & B_k \\ & & & & D_{k+1} \end{bmatrix} \in \mathbb{R}^{(ku+r) \times (ku+r)},$$

$$\tilde{D}_1 \in \mathbb{R}^{r \times r}, \quad \tilde{B}_1 \in \mathbb{R}^{r \times u};$$

$$D_i \in \mathbb{R}^{u \times u}, \quad B_i \in \mathbb{R}^{u \times u} \text{ for } i = 2, \dots, k+1.$$

$$X = \begin{bmatrix} \tilde{X}_1 \\ X_2 \\ \vdots \\ X_{k+1} \end{bmatrix} \in \mathbb{R}^{(ku+r) \times u} \quad Y = \begin{bmatrix} \tilde{Y}_1 \\ Y_2 \\ \vdots \\ Y_{k+1} \end{bmatrix} \in \mathbb{R}^{(ku+r) \times u} \quad Y^T = [\tilde{Y}_1^T \quad Y_2^T \quad \cdots \quad Y_{k+1}^T] \in \mathbb{R}^{u \times (ku+r)}$$

$$\begin{aligned}
&\tilde{X}_1 \in \mathbb{R}^{r \times u} \\
&\tilde{Y}_1^T \in \mathbb{R}^{u \times r} \\
&X_i, Y_i^T \in \mathbb{R}^{u \times u} \text{ for } i = 2, \dots, k+1 \\
&X := U^{-1}E_{k+1} \\
&\tilde{Y}_1^T := (D_1X_1)^+ = (D_1X_1)^T(D_1X_1(D_1X_1)^T)^{-1} \\
&\text{triu}(D_1X_1\tilde{Y}_1^T) = I_{r \times r} \\
&Y_i^T := (D_iX_i)^{-1} \text{ for } i = 2, \dots, k+1
\end{aligned}$$

## 17 28.06.2022 Meeting formal

### 17.1 Notation

This notes use Householder notation and start indexing at one.

Scalars, lower case Greek letters:  $\alpha, \beta, \gamma, \chi$ , *etc.*

Vectors, lower case Roman letter:  $a, b, c, x$ , *etc.*

Matrix, upper case Roman letter:  $A, B, C, X$ , *etc.*

### 17.2 Repersentation

#### Matrix & Vector

$$x = \begin{bmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{bmatrix} \in \mathbb{R}^n$$
$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{1n} \end{bmatrix}$$
$$= [a_1 \mid a_2 \mid \cdots \mid a_n] \in \mathbb{R}^{m \times n}$$

#### Block Matrix & Block Vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{kn}, x_i \in \mathbb{R}^k$$
$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{1n} \end{bmatrix}$$
$$= [A_1 \mid A_2 \mid \cdots \mid A_n] \in \mathbb{R}^{lm \times kn}, A_{ij} \in \mathbb{R}^{l \times k}$$

### 17.3 Content

Consider an upper banded matrix  $U \in \mathbb{R}^{n \times n}$ ,

$$U = \begin{bmatrix} \alpha_1 & \beta_1^1 & \beta_1^2 & \dots & \beta_1^u & & \\ & \alpha_2 & \beta_2^1 & \beta_2^2 & \dots & \beta_2^u & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & \beta_{n-u}^u & \\ & & & & & \vdots & \\ & & & & & \beta_{n-2}^2 & \\ & & & & & \beta_{n-1}^1 & \\ & & & & & & \alpha_n \end{bmatrix}. \quad (54)$$

Then try to show that its inverse can be represented as the upper triangular part of an outer product, *i.e.*,

$$U^{-1} = \text{triu}(XY^T) \quad (55)$$

Let  $X, Y \in \mathbb{R}^{n \times u}$ , consider the following cases of  $n$ ,

#### Case 1, $n = ku$

then (54) can be represented as a block upper bidiagonal matrix,

$$\begin{bmatrix} \xrightarrow{r} & \xleftarrow{u} & \dots & \xleftarrow{u} \\ \begin{matrix} d_1 & \dots & b_1^{r-1} \\ & \ddots & \vdots \\ & & d_r \end{matrix} & \begin{matrix} b_1^r & \dots & b_1^u \\ \vdots & \ddots & \vdots \\ b_r^1 & \dots & b_r^u \end{matrix} & \dots & \dots \\ d_{1+r} & \dots & b_{1+r}^{u-1} & b_{r+1}^u & \dots & \\ & \ddots & \vdots & \vdots & \ddots & \\ & & d_{u+r} & b_{u+r}^1 & \dots & b_{u+r}^u \\ & & & \ddots & \ddots & \ddots \\ & & & & d_{n-u+1} & \dots & b_{n-u+1}^u \\ & & & & & \ddots & \vdots \\ & & & & & & d_n \end{bmatrix}$$



$$\begin{aligned}
U &= \begin{bmatrix} \xrightarrow{u} & \xrightarrow{u} & \dots & \xrightarrow{u} \\ \alpha_1 \dots \beta_1^{u-1} & \beta_1^u & & \\ \vdots & \vdots & \ddots & \\ \alpha_u & \beta_u^1 \dots \beta_u^u & & \\ & \alpha_{u+1} \dots \beta_{u+1}^{u-1} & \beta_{u+1}^u & \\ & \vdots & \vdots & \ddots \\ & \alpha_{2u} & \beta_{2u}^1 \dots \beta_{2u}^u & \\ & & \ddots & \\ & & & \alpha_{(k-1)u+1} \dots \beta_{(k-1)u+1}^{u-1} \\ & & & \vdots \\ & & & \alpha_{ku} \end{bmatrix} \\
&= \begin{bmatrix} D_1 & B_1 & & & \\ & D_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & D_{k-1} & B_{k-1} \\ & & & & D_k \end{bmatrix} \in \mathbb{R}^{ku \times ku}, D_i, B_i \in \mathbb{R}^{u \times u} \text{ for } i = 1, 2, \dots, k. \quad (56)
\end{aligned}$$

Represent the  $n$ -dimensional identity matrix  $I_{n \times n}$  in block form,

$$I_{n \times n} = I_{ku \times ku} = \text{diag}(\underbrace{I_{u \times u}, I_{u \times u}, \dots, I_{u \times u}}_k) = [E_1 \mid E_2 \mid \dots \mid E_k].$$

then, its block columns  $E_i$  for  $i = 1, 2, \dots, k$  satisfies

$$\begin{aligned}
E_i &\in \mathbb{R}^{ku \times u} \\
E_i^T E_j &= \begin{cases} I_{u \times u}, & \text{if } i = j; \\ 0_{u \times u}, & \text{otherwise.} \end{cases}
\end{aligned}$$

To find  $X$  and  $Y$  such that (55) holds, let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{k-1} \\ X_k \end{bmatrix} \in \mathbb{R}^{ku \times u}, Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{k-1} \\ Y_k \end{bmatrix} \in \mathbb{R}^{ku \times u}, X_i, Y_i \in \mathbb{R}^{u \times u} \text{ for } i = 1, 2, \dots, k,$$

and let  $X_k$  and  $Y_k^T$  be upper triangular matrix, then  $\text{triu}(X_k Y_k^T) = X_k Y_k^T$ .  
For the outer product form,

$$\begin{aligned}
Y^T &= [ Y_1^T \mid Y_2^T \mid \cdots \mid Y_{k-1}^T \mid Y_k^T ] \in \mathbb{R}^{u \times ku} \\
XY^T &= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{k-1} \\ X_k \end{bmatrix} [ Y_1^T \mid Y_2^T \mid \cdots \mid Y_{k-1}^T \mid Y_k^T ] \\
&= \begin{bmatrix} X_1 Y_1^T & X_1 Y_2^T & \cdots & X_1 Y_{k-1}^T & X_1 Y_k^T \\ X_2 Y_1^T & X_2 Y_2^T & \cdots & X_2 Y_{k-1}^T & X_2 Y_k^T \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{k-1} Y_1^T & X_{k-1} Y_2^T & \cdots & X_{k-1} Y_{k-1}^T & X_{k-1} Y_k^T \\ X_k Y_1^T & X_k Y_2^T & \cdots & X_k Y_{k-1}^T & X_k Y_k^T \end{bmatrix} \\
&= X [ Y_1^T \mid Y_2^T \mid \cdots \mid Y_{k-1}^T \mid Y_k^T ] \\
&= [ XY_1^T \mid XY_2^T \mid \cdots \mid XY_{k-1}^T \mid XY_k^T ] \in \mathbb{R}^{ku \times ku},
\end{aligned}$$

and its the upper triangular component,

$$\begin{aligned}
\text{triu}(XY^T) &= \begin{bmatrix} \text{triu}(X_1 Y_1^T) & X_1 Y_2^T & \cdots & X_1 Y_{k-1}^T & X_1 Y_k^T \\ 0_{u \times u} & \text{triu}(X_2 Y_2^T) & \cdots & X_2 Y_{k-1}^T & X_2 Y_k^T \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_{u \times u} & 0_{u \times u} & \cdots & \text{triu}(X_{k-1} Y_{k-1}^T) & X_{k-1} Y_k^T \\ 0_{u \times u} & 0_{u \times u} & \cdots & 0_{u \times u} & X_k Y_k^T \end{bmatrix} \\
&= \begin{bmatrix} \text{triu}(X_1 Y_1^T) & X_1 Y_2^T & \cdots & X_{1:k-2} Y_{k-1}^T & X_1 Y_k^T \\ 0_{(k-2)u \times u} & \text{triu}(X_2 Y_2^T) & \cdots & \text{triu}(X_{k-1} Y_{k-1}^T) & X_2 Y_k^T \\ 0_{u \times u} & 0_{(k-2)u \times u} & \cdots & 0_{u \times u} & X_k Y_k^T \end{bmatrix} .zz \quad (57)
\end{aligned}$$

From (55),

$$I_{n \times n} = U \text{triu}(XY^T). \quad (58)$$

For  $X$  and  $Y_k$ , from (57), the  $k$ -th block column of (58) can be written as,

$$\begin{aligned}
E_k &= U \text{triu}(XY^T) E_k \\
E_k &= UXY_k^T \quad (59)
\end{aligned}$$

Define that

$$X := U^{-1} E_k; \quad (60)$$

$$Y_k^T := I_{u \times u}, \quad (61)$$

which can be shown that (59) holds for the defined (60) and (61) through

$$\begin{aligned}
UXY_k^T &= UU^{-1} E_k I_{u \times u} \\
&= I_{n \times n} E_k I_{u \times u} \\
&= E_k I_{u \times u} \\
&= E_k, \text{ as required.}
\end{aligned}$$

Then, for  $Y_{k-1}$ , forming the  $(k-1)$ -th block column of (58) from (57) first,

$$\begin{aligned} E_{k-1} &= U \text{triu}(XY^T)E_{k-1} \\ &= U \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \text{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{u \times u} \end{bmatrix}. \end{aligned}$$

Focusing on the  $(k-1)$ -th block row of  $(k-1)$ -th block column,

$$\begin{aligned} E_{k-1}^T E_{k-1} &= E_{k-1}^T U \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \text{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{u \times u} \end{bmatrix} \\ I_{u \times u} &= \left[ \begin{array}{c|c|c} 0_{u \times (k-2)u} & D_{k-1} & B_{k-1} \end{array} \right] \begin{bmatrix} X_{1:k-2}Y_{k-1}^T \\ \text{triu}(X_{k-1}Y_{k-1}^T) \\ 0_{u \times u} \end{bmatrix} \\ I_{u \times u} &= D_{k-1} \text{triu}(X_{k-1}Y_{k-1}^T). \end{aligned} \tag{62}$$

Define that

$$Y_{k-1}^T := (D_{k-1}X_{k-1})^{-1}, \tag{63}$$

and notice that  $D_{k-1} \in \mathbb{R}^{u \times u}$  is a upper triangular, therefore

$$\begin{aligned} D_{k-1} &= \text{triu}(D_{k-1}); \\ D_{k-1}^{-1} &= \text{triu}(D_{k-1}^{-1}). \end{aligned} \tag{64}$$

With the fact (64), (62) holds for the defined (63) as,

$$\begin{aligned} D_{k-1} \text{triu}(X_{k-1}Y_{k-1}^T) &= D_{k-1} \text{triu}(X_{k-1}(D_{k-1}X_{k-1})^{-1}) \\ &= D_{k-1} \text{triu}(X_{k-1}X_{k-1}^{-1}D_{k-1}^{-1}) \\ &= D_{k-1} \text{triu}(D_{k-1}^{-1}) \\ &= D_{k-1}D_{k-1}^{-1} \\ &= I_{u \times u}, \text{ as required.} \end{aligned}$$

Similar ideas can be use to redefine  $Y_k^T$  in (61), from (60),

$$\begin{aligned} E_k &= UX \\ E_k^T E_k &= E_k^T UX \\ I_{u \times u} &= E_k^T UX \\ &= \left[ \begin{array}{c|c} 0_{u \times (k-1)u} & D_k \end{array} \right] \begin{bmatrix} X_{1:k-1} \\ X_k \end{bmatrix} \\ &= D_k X_k \\ I_{u \times u} &= (D_k X_k)^{-1}, \end{aligned}$$

following that

$$\begin{aligned}
Y_k^{\text{T}} &:= I_{u \times u} \\
&:= (D_k X_k)^{-1}.
\end{aligned}$$

In summary, for  $U \in \mathbb{R}^{ku \times ku}$ , (55) holds the following definitions of  $X$  and  $Y$ :

$$\begin{aligned}
X &:= U^{-1} E_k \\
Y_i^{\text{T}} &:= (D_i X_i)^{-1}
\end{aligned}$$

[?]