Numerical Linear Algebra I

Titus Pinta

September 4, 2018

Contents I

Contents II

Matrix Multiplication

Matrix Multiplication:

$$A \in M^{mxn}(\mathbb{K}),\ B \in M^{nxp}(\mathbb{K}),\ AB \in M^{mxp}(\mathbb{K})$$

Matrix Multiplication

Matrix Multiplication:

$$A \in M^{mxn}(\mathbb{K}),\ B \in M^{nxp}(\mathbb{K}),\ AB \in M^{mxp}(\mathbb{K})$$

Matrix Block Multiplication:

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix} = \begin{bmatrix} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} & A_{1,1}B_{1,2} + A_{1,2}B_{2,2} \\ A_{2,1}B_{1,1} + A_{2,2}B_{2,1} & A_{2,1}B_{1,2} + A_{2,2}B_{2,2} \end{bmatrix}$$

Matrix Multiplication

Matrix Multiplication:

$$A \in M^{mxn}(\mathbb{K}),\ B \in M^{nxp}(\mathbb{K}),\ AB \in M^{mxp}(\mathbb{K})$$

Matrix Block Multiplication:

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix} = \begin{bmatrix} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} & A_{1,1}B_{1,2} + A_{1,2}B_{2,2} \\ A_{2,1}B_{1,1} + A_{2,2}B_{2,1} & A_{2,1}B_{1,2} + A_{2,2}B_{2,2} \end{bmatrix}$$

Naive algorithm:
Additions: $mp(n-1)$
Multiplications: mnp
Operations count $= \mathcal{O}(mnp) = \mathcal{O}(n^3)$ for square matrices

Strassen's algorithm

$$M_1 = (A_{1,1} + A_{2,2})(B_{1,1} + B_{2,2})$$
$$M_2 = (A_{2,1} + A_{2,2})B_{1,1}$$
$$M_3 = A_{1,1}(B_{1,2} - B_{2,2})$$
$$M_4 = A_{2,2}(B_{2,1} - B_{1,1})$$
$$M_5 = (A_{1,1} + A_{1,2})B_{2,2}$$
$$M_6 = (A_{2,1} - A_{1,1})(B_{1,1} - B_{1,2})$$
$$M_7 = (A_{1,2} - A_{2,2})(B_{2,1} - B_{2,2})$$

$$AB = \begin{bmatrix} M_1 + M_4 - M_5 + M_7 & M_3 + M_5 \\ M_2 + M_4 & M_1 - M_2 + M_3 + M_6 \end{bmatrix}$$

Strassen's algorithm

$$
\begin{aligned}
M_1 &= (A_{1,1} + A_{2,2})(B_{1,1} + B_{2,2}) \\
M_2 &= (A_{2,1} + A_{2,2})B_{1,1} \\
M_3 &= A_{1,1}(B_{1,2} - B_{2,2}) \\
M_4 &= A_{2,2}(B_{2,1} - B_{1,1}) \\
M_5 &= (A_{1,1} + A_{1,2})B_{2,2} \\
M_6 &= (A_{2,1} - A_{1,1})(B_{1,1} - B_{1,2}) \\
M_7 &= (A_{1,2} - A_{2,2})(B_{2,1} - B_{2,2})
\end{aligned}
$$

$$
AB = \begin{bmatrix} M_1 + M_4 - M_5 + M_7 & M_3 + M_5 \\ M_2 + M_4 & M_1 - M_2 + M_3 + M_6 \end{bmatrix}
$$

Complexity: $\mathcal{O}(7^{\log_2 n}) = \mathcal{O}(n^{\log_2 7})$

Condition Numbers

Let X, Y be normed vector spaces, and let $f : X \to Y$ be a function (problem)

Condition Numbers

Let $X, \ Y$ be normed vector spaces, and let $f : X \to Y$ be a function (problem)

Absolute condition number:

$$\hat{\kappa} = \lim_{\delta \to 0} \sup_{\|\delta x\| < \delta} \frac{\|f(x + \delta x) - f(x)\|}{\|\delta x\|}$$

Condition Numbers

Let $X,\ Y$ be normed vector spaces, and let $f : X \to Y$ be a function (problem)

Absolute condition number:

$$\hat{\kappa} = \lim_{\delta \to 0} \sup_{\|\delta x\| < \delta} \frac{\|f(x+\delta x) - f(x)\|}{\|\delta x\|}$$

Relative condition number:

$$\kappa = \lim_{\delta \to 0} \sup_{\|\delta x\| < \delta} \frac{\|f(x+\delta x) - f(x)\|}{\|f(x)\|} \frac{\|x\|}{\|\delta x\|}$$

Stability

Let f be a problem and $\hat{f}$ a computer algorithm for f

Stability

Let f be a problem and $\hat{f}$ a computer algorithm for f
Accuracy:

$$\frac{\|f(x) - \hat{f}(x)\|}{\|f(x)\|} = \mathcal{O}(\epsilon)$$

Stability

Let f be a problem and $\hat{f}$ a computer algorithm for f
Accuracy:

$$\frac{\|f(x) - \hat{f}(x)\|}{\|f(x)\|} = \mathcal{O}(\epsilon)$$

Stability:

$$\exists \hat{x}, \frac{\|x - \hat{x}\|}{\|x\|} = \mathcal{O}(\epsilon)$$

$$\frac{\|\hat{f}(x) - f(\hat{x})\|}{\|f(\hat{x})\|} = \mathcal{O}(\epsilon)$$

Backward Stability

$$\exists \hat{x}, \frac{\|x - \hat{x}\|}{\|x\|} = \mathcal{O}(\epsilon)$$

$$\hat{f}(x) = f(\hat{x})$$

Backward Stability

$$\exists \hat{x}, \frac{\|x - \hat{x}\|}{\|x\|} = \mathcal{O}(\epsilon)$$

$$\hat{f}(x) = f(\hat{x})$$

Backward stability on a well-conditioned problem: Let f be backward stable and well-conditioned $\kappa \approx 1$

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq C\epsilon$$

$$\frac{\|f(\hat{x}) - f(x)\|}{\|f(x)\|} \frac{\|x\|}{\|x - \hat{x}\|} \leq \kappa$$

$$\frac{\|f(\hat{x}) - f(x)\|}{\|f(x)\|} = \mathcal{O}(\kappa\epsilon) \Rightarrow \frac{\|\hat{f}(x) - f(x)\|}{\|f(x)\|} = \mathcal{O}(\epsilon)$$

BLAS

Level 1: Vector-Vector Operations

$$y \leftarrow \alpha x + y$$

BLAS

Level 1: Vector-Vector Operations

$$y \leftarrow \alpha x + y$$

Level 2: Matrix-Vector Operations

$$y \leftarrow \alpha A x + y$$

BLAS

Level 1: Vector-Vector Operations

$$y \leftarrow \alpha x + y$$

Level 2: Matrix-Vector Operations

$$y \leftarrow \alpha A x + y$$

Level 3: Matrix-Matrix Operations

$$C \leftarrow \alpha AB + \beta C$$

Condition of linear systems

Let $A \in M^n(\mathbb{K})$ with $\det A \neq 0$
The problem: Find x so $Ax = b$
Let $\| \cdot \|$ be a norm on $\mathbb{K}^n$ and let $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
Let A be fixed

Condition of linear systems

Let $A \in M^n(\mathbb{K})$ with $\det A \neq 0$
The problem: Find x so $Ax = b$
Let $\|\cdot\|$ be a norm on $\mathbb{K}^n$ and let $\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$
Let A be fixed

$$Ax = b \Rightarrow x = A^{-1}b$$

$$A\hat{x} = b + \delta b \Rightarrow \hat{x} = A^{-1}(b + \delta b)$$

Condition of linear systems

Let $A \in M^n(\mathbb{K})$ with $\det A \neq 0$
The problem: Find x so $Ax = b$
Let $\|\cdot\|$ be a norm on $\mathbb{K}^n$ and let $\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$
Let A be fixed

$$Ax = b \Rightarrow x = A^{-1}b$$

$$A\hat{x} = b + \delta b \Rightarrow \hat{x} = A^{-1}(b + \delta b)$$

$$\|x - \hat{x}\| \leq \|A^{-1}\|\|\delta b\|$$

$$\|b\| \leq \|A\|\|x\|$$

Condition of linear systems

Let $A \in M^n(\mathbb{K})$ with $\det A \neq 0$
The problem: Find x so $Ax = b$
Let $\| \cdot \|$ be a norm on $\mathbb{K}^n$ and let $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
Let A be fixed

$$Ax = b \Rightarrow x = A^{-1} b$$

$$A\hat{x} = b + \delta b \Rightarrow \hat{x} = A^{-1}(b + \delta b)$$

$$\|x - \hat{x}\| \leq \|A^{-1}\| \|\delta b\|$$

$$\|b\| \leq \|A\| \|x\|$$

$$\frac{\|x - \hat{x}\|}{\|x\|} \frac{\|b\|}{\|\delta b\|} \leq \|A\| \|A^{-1}\|$$

Condition of linear systems

Let $x \in \mathbb{K}^n$ with $\|Ax\| = \|A\|\|x\|$ and $\delta b \in \mathbb{K}^n$ with $\|A^{-1}\delta b\| = \|A^{-1}\|\|\delta b\|$

Condition of linear systems

Let $x \in \mathbb{K}^n$ with $\|Ax\| = \|A\|\|x\|$ and $\delta b \in \mathbb{K}^n$ with $\|A^{-1}\delta b\| = \|A^{-1}\|\|\delta b\|$

$$\frac{\|x - \hat{x}\|}{\|x\|} \frac{\|b\|}{\|\delta b\|} = \frac{\|A^{-1}\delta b\|}{\|x\|} \frac{\|Ax\|}{\|\delta b\|} = \|A\|\|A^{-1}\|$$

Condition of linear systems

Let $x \in \mathbb{K}^n$ with $\|Ax\| = \|A\|\|x\|$ and $\delta b \in \mathbb{K}^n$ with $\|A^{-1}\delta b\| = \|A^{-1}\|\|\delta b\|$

$$\frac{\|x - \hat{x}\|}{\|x\|} \frac{\|b\|}{\|\delta b\|} = \frac{\|A^{-1}\delta b\|}{\|x\|} \frac{\|Ax\|}{\|\delta b\|} = \|A\|\|A^{-1}\|$$

When A is fixed:

$$\kappa = \|A\|\|A^{-1}\|$$

Condition of linear systems

Suppose b is fixed

$$Ax = b, \; (A + \delta A)\hat{x} = b$$

Condition of linear systems

Suppose b is fixed

$$Ax = b, \; (A + \delta A)\hat{x} = b$$

$$A\hat{x} + \delta A\hat{x} = Ax \Rightarrow \hat{x} - x = -A^{-1}\delta A\hat{x}$$

Condition of linear systems

Suppose b is fixed

$$Ax = b, \ (A + \delta A)\hat{x} = b$$

$$A\hat{x} + \delta A\hat{x} = Ax \Rightarrow \hat{x} - x = -A^{-1}\delta A\hat{x}$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|} \leq \|A^{-1}\|\|\delta A\|\frac{\|A\|}{\|A\|} \Rightarrow \frac{\|\hat{x} - x\|}{\|\hat{x}\|}\frac{\|A\|}{\|\delta A\|} \leq \|A\|\|A^{-1}\|$$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

$$x - \hat{x} = \beta A^{-1}\hat{x} \Rightarrow \|x - \hat{x}\| = |\beta|\|A^{-1}\|\|\hat{x}\|$$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

$$x - \hat{x} = \beta A^{-1}\hat{x} \Rightarrow \|x - \hat{x}\| = |\beta|\|A^{-1}\|\|\hat{x}\|$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|}\frac{\|A\|}{\|\delta A\|} = \frac{|\beta|\|A^{-1}\|\|\hat{x}\|}{\|\hat{x}\|}\frac{\|A\|}{|\beta|} = \|A\|\|A^{-1}\|$$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

$$x - \hat{x} = \beta A^{-1}\hat{x} \Rightarrow \|x - \hat{x}\| = |\beta|\|A^{-1}\|\|\hat{x}\|$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|}\frac{\|A\|}{\|\delta A\|} = \frac{|\beta|\|A^{-1}\|\|\hat{x}\|}{\|\hat{x}\|}\frac{\|A\|}{|\beta|} = \|A\|\|A^{-1}\|$$

Condition number of a matrix:

$$\kappa(A) = \|A\|\|A^{-1}\|$$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

$$x - \hat{x} = \beta A^{-1}\hat{x} \Rightarrow \|x - \hat{x}\| = |\beta|\|A^{-1}\|\|\hat{x}\|$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|}\frac{\|A\|}{\|\delta A\|} = \frac{|\beta|\|A^{-1}\|\|\hat{x}\|}{\|\hat{x}\|}\frac{\|A\|}{|\beta|} = \|A\|\|A^{-1}\|$$

Condition number of a matrix:

$$\kappa(A) = \|A\|\|A^{-1}\|$$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \kappa(A) \geq 1$$

Condition of linear systems

Let $\hat{x} \in \mathbb{K}^n$ with $\|A^{-1}\hat{x}\| = \|A^{-1}\|\|\hat{x}\|$ and $\beta \in \mathbb{K}$, $\delta A = \beta I$

$$x - \hat{x} = \beta A^{-1}\hat{x} \Rightarrow \|x - \hat{x}\| = |\beta|\|A^{-1}\|\|\hat{x}\|$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|}\frac{\|A\|}{\|\delta A\|} = \frac{|\beta|\|A^{-1}\|\|\hat{x}\|}{\|\hat{x}\|}\frac{\|A\|}{|\beta|} = \|A\|\|A^{-1}\|$$

Condition number of a matrix:

$$\kappa(A) = \|A\|\|A^{-1}\|$$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \kappa(A) \geq 1$$

Remark: the larger κ, the more ill-conditioned the linear system is

Nearest singular matrix

Let A be non-singular and B singular $\Rightarrow Bx = 0$ for some x with $\|x\| = 1$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \|A\| \geq \|A^{-1}\|^{-1}$$

Nearest singular matrix

Let A be non-singular and B singular $\Rightarrow Bx = 0$ for some x with $\|x\| = 1$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \|A\| \geq \|A^{-1}\|^{-1}$$

$$\|A - B\| = \|A - B\|\|x\| \geq \|(A - B)x\| = \|Ax - Bx\| = \|Ax\|$$

Nearest singular matrix

Let A be non-singular and B singular $\Rightarrow$ $Bx = 0$ for some x with $\|x\| = 1$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \|A\| \geq \|A^{-1}\|^{-1}$$

$$\|A - B\| = \|A - B\|\|x\| \geq \|(A - B)x\| = \|Ax - Bx\| = \|Ax\|$$

$$1 = \|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\|\|Ax\| \Rightarrow \|Ax\| \geq \|A^{-1}\|^{-1}$$

Nearest singular matrix

Let A be non-singular and B singular $\Rightarrow Bx = 0$ for some x with $\|x\| = 1$

$$1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \Rightarrow \|A\| \geq \|A^{-1}\|^{-1}$$

$$\|A - B\| = \|A - B\|\|x\| \geq \|(A - B)x\| = \|Ax - Bx\| = \|Ax\|$$

$$1 = \|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\|\|Ax\| \Rightarrow \|Ax\| \geq \|A^{-1}\|^{-1}$$

$$\|A - B\| \geq \|A^{-1}\|^{-1}$$

Nearest singular matrix

$$\|y^T\|_D = \sup_{x \neq 0} \frac{y^T x}{\|x\|}$$

Observe:

$$\|A\| = \sup_{y \neq 0} \frac{\|y^T A\|_D}{\|y\|_D}$$

Nearest singular matrix

$$\|y^T\|_D = \sup_{x \neq 0} \frac{y^T x}{\|x\|}$$

Observe:

$$\|A\| = \sup_{y \neq 0} \frac{\|y^T A\|_D}{\|y\|_D}$$

$\exists x,\ y \in \mathbb{K}^n$ with $\|x\| = \|y\|_D = 1$

$$y^T A^{-1} x = \|A^{-1}\|$$

Nearest singular matrix

$$\delta A = -\|A^{-1}\|^{-1} x y^T$$

Nearest singular matrix

$$\delta A = -\|A^{-1}\|^{-1} x y^T$$

$$x \neq 0 \Rightarrow A^{-1} x \neq 0$$

$$(A+\delta A)A^{-1}x = AA^{-1}x - \|A^{-1}\|^{-1} x y^T A^{-1} x = x - \|A^{-1}\|^{-1}\|A^{-1}\| x = 0$$

$$\det(A + \delta A) = 0$$

Nearest singular matrix

$$\delta A = -\|A^{-1}\|^{-1} x y^T$$

$$x \neq 0 \Rightarrow A^{-1}x \neq 0$$

$$(A+\delta A)A^{-1}x = AA^{-1}x - \|A^{-1}\|^{-1} x y^T A^{-1} x = x - \|A^{-1}\|^{-1}\|A^{-1}\| x = 0$$

$$\det(A+\delta A) = 0$$

Let $\|v\| = 1$ and $y^T v = 1$

$$1 = \|x\| = \|xy^T v\| \leq \|xy^T\|\|v\| = \|xy^T\| \leq \|x\|\|y^T\|_D = 1$$

$$\|\delta A\| = \|A^{-1}\|^{-1}$$

Nearest singular matrix

$$\delta A = -\|A^{-1}\|^{-1} x y^T$$

$$x \neq 0 \Rightarrow A^{-1}x \neq 0$$

$$(A+\delta A)A^{-1}x = AA^{-1}x - \|A^{-1}\|^{-1} x y^T A^{-1} x = x - \|A^{-1}\|^{-1}\|A^{-1}\| x = 0$$

$$\det(A + \delta A) = 0$$

Let $\|v\| = 1$ and $y^T v = 1$

$$1 = \|x\| = \|xy^T v\| \leq \|xy^T\|\|v\| = \|xy^T\| \leq \|x\|\|y^T\|_D = 1$$

$$\|\delta A\| = \|A^{-1}\|^{-1}$$

$$\inf_{\det B = 0} \frac{\|A - B\|}{\|A\|} = \frac{1}{\kappa(A)}$$

LU Factorization:

$$\begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ A_{n,1} & \dots & A_{n,n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ L_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ L_{n,1} & L_{n,2} & \dots & 1 \end{bmatrix} \begin{bmatrix} U_{1,1} & U_{1,2} & \dots & U_{1,n} \\ 0 & U_{2,2} & \dots & U_{2,n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & U_{n,n} \end{bmatrix}$$

LU Factorization:

$$\begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ A_{n,1} & \dots & A_{n,n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ L_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ L_{n,1} & L_{n,2} & \dots & 1 \end{bmatrix} \begin{bmatrix} U_{1,1} & U_{1,2} & \dots & U_{1,n} \\ 0 & U_{2,2} & \dots & U_{2,n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & U_{n,n} \end{bmatrix}$$

Gaussian elimination:

$$U = \widehat{L_n}\widehat{L_{n-1}} \dots \widehat{L_1} A$$

$$\widehat{L_k} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 & \\ 0 & \dots & -\frac{a_{k+1,k}}{a_{k,k}} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -\frac{a_{n,k}}{a_{k,k}} & 0 & \dots & 1 \end{bmatrix}$$

Gaussian elimination

$$A = \widehat{L_1^{-1}} \dots \widehat{L_n^{-1}} U$$

$$\widehat{L_k^{-1}} = -\widehat{L_k} =: L_k$$

$$A = L_1 \dots L_n U = LU$$

Gaussian elimination

$$A = \widehat{L_1^{-1}} \dots \widehat{L_n^{-1}} U$$

$$\widehat{L_k^{-1}} = -\widehat{L_k} =: L_k$$

$$A = L_1 \dots L_n U = LU$$

Linear systems:

$$Ax = LUx = b$$

$$Ly = x \Rightarrow y = L^{-1}x$$

$$Ux = y \Rightarrow x = U^{-1}y$$

Gaussian elimination

$$A = \widehat{L_1^{-1}} \dots \widehat{L_n^{-1}} U$$

$$\widehat{L_k^{-1}} = -\widehat{L_k} =: L_k$$

$$A = L_1 \dots L_n U = LU$$

Linear systems:

$$Ax = LUx = b$$

$$Ly = x \Rightarrow y = L^{-1}x$$

$$Ux = y \Rightarrow x = U^{-1}y$$

$$Ax = b \Rightarrow \widehat{L_1} Ax = \widehat{L_1} b \Rightarrow \widehat{L_n}\widehat{L_{n-1}} \dots \widehat{L_1} Ax = \widehat{L_n}\widehat{L_{n-1}} \dots \widehat{L_1} b$$

$$Ux = \widehat{L_n}\widehat{L_{n-1}} \dots \widehat{L_1} b$$

Backsubstitution

$$Ux = b$$

$$u_{1,1}x_1 + u_{1,2}x_2 + \ldots + u_{1,n}x_n = b_1$$
$$u_{2,2}x_2 + \ldots + u_{2,n}x_n = b_1$$
$$\ldots$$
$$u_{n,n}x_n = b_n$$

Backsubstitution

$$Ux = b$$

$$u_{1,1}x_1 + u_{1,2}x_2 + \ldots + u_{1,n}x_n = b_1$$

$$u_{2,2}x_2 + \ldots + u_{2,n}x_n = b_1$$

$$\ldots$$

$$u_{n,n}x_n = b_n$$

$$x_n = \frac{b_n}{u_{n,n}}$$

$$x_{n-1} = \frac{b_{n-1} - u_{n-1,n}x_n}{u_{n-1,n-1}}$$

$$\ldots$$

$$x_1 = \frac{b_1 - (u_{1,n}x_n + u_{1,n-1}x_{n-1} + \ldots + u_{1,2}x_2)}{u_{1,1}}$$

Backsubstitution

Operations count:
Consider $\bar{A} = [A; b] \in M^{n,n+1}(\mathbb{K})$

Backsubstitution

Operations count:
Consider $\bar{A} = [A; b] \in M^{n,n+1}(\mathbb{K})$
LU Factorization:
For one 0 $\Rightarrow$ $n + 1 - k$ additions and multiplications
There are $\frac{n(n-1)}{2}$ below diagonal

Backsubstitution

Operations count:
Consider $\bar{A} = [A; b] \in M^{n,n+1}(\mathbb{K})$
LU Factorization:
For one 0 $\Rightarrow$ $n + 1 - k$ additions and multiplications
There are $\frac{n(n-1)}{2}$ below diagonal
Backsubstitution:
For x_{n-k} we need $k + 1$ multiplications and k additions

Operations count:
Consider $\bar{A} = [A; b] \in M^{n,n+1}(\mathbb{K})$
LU Factorization:
For one 0 $\Rightarrow$ $n + 1 - k$ additions and multiplications
There are $\frac{n(n-1)}{2}$ below diagonal
Backsubstitution:
For x_{n-k} we need $k + 1$ multiplications and k additions
Total:
Additions: $\frac{(n-1)n(n+1)+n(n+1)}{2}$
Multiplications: $\frac{(n-1)n(n+1)+n(n+3)}{2}$
Operations count $= \mathcal{O}(n^3)$

Stability of LU

$Ux = b$ is backwards stable

Stability of LU

$Ux = b$ is backwards stable
For x_n:

$$u_{n,n}x_n = b_n \Rightarrow \widehat{x_n} = \frac{b_n}{u_{n,n}}(1 + \epsilon_1)$$

Stability of LU

$Ux = b$ is backwards stable
For x_n:

$$u_{n,n}x_n = b_n \Rightarrow \widehat{x_n} = \frac{b_n}{u_{n,n}}(1+\epsilon_1)$$

$$\epsilon_1 = \mathcal{O}(\epsilon) \Rightarrow \epsilon_2 = \frac{\epsilon_1}{1+\epsilon_1} = \epsilon_1 - \epsilon_1^2 + \epsilon_1^3 - \ldots = O(\epsilon)$$

Stability of LU

$Ux = b$ is backwards stable
For x_n:

$$u_{n,n}x_n = b_n \Rightarrow \widehat{x_n} = \frac{b_n}{u_{n,n}}(1+\epsilon_1)$$

$$\epsilon_1 = \mathcal{O}(\epsilon) \Rightarrow \epsilon_2 = \frac{\epsilon_1}{1+\epsilon_1} = \epsilon_1 - \epsilon_1^2 + \epsilon_1^3 - \ldots = O(\epsilon)$$

$$\epsilon_2 = \frac{\epsilon_1}{1+\epsilon_1} \Rightarrow \epsilon_1 = \frac{-\epsilon_2}{1+\epsilon_2}$$

$$\widehat{x_n} = \frac{b_n}{u_{n,n}} \frac{1+\epsilon_2-\epsilon_2}{1+\epsilon_2}$$

Stability of LU

$Ux = b$ is backwards stable
For x_n:

$$u_{n,n}x_n = b_n \Rightarrow \widehat{x_n} = \frac{b_n}{u_{n,n}}(1+\epsilon_1)$$

$$\epsilon_1 = \mathcal{O}(\epsilon) \Rightarrow \epsilon_2 = \frac{\epsilon_1}{1+\epsilon_1} = \epsilon_1 - \epsilon_1^2 + \epsilon_1^3 - \ldots = O(\epsilon)$$

$$\epsilon_2 = \frac{\epsilon_1}{1+\epsilon_1} \Rightarrow \epsilon_1 = \frac{-\epsilon_2}{1+\epsilon_2}$$

$$\widehat{x_n} = \frac{b_n}{u_{n,n}}\frac{1+\epsilon_2-\epsilon_2}{1+\epsilon_2}$$

$$(u_{n,n} + \epsilon_2 u_{n,n})\widehat{x_n} = b_n, \;\; \epsilon_2 = O(\epsilon)$$

For x_{n-k}:

$$x_{n-k} = \frac{b_{n-k} - \left(u_{n-k,n}x_n + u_{n-k,n-1}x_{n-1} \ldots u_{n-k,n-k+1}x_{n-k+1}\right)}{u_{n-k,n-k}}$$

For x_{n-k}:

$$x_{n-k} = \frac{b_{n-k} - (u_{n-k,n}x_n + u_{n-k,n-1}x_{n-1} \ldots u_{n-k,n-k+1}x_{n-k+1})}{u_{n-k,n-k}}$$

$$\widehat{x_{n-k}} = \frac{\begin{array}{c} b_{n-k} - (u_{n-k,n}\widehat{x_n}(1+\epsilon) + u_{n-k,n-1}\widehat{x_{n-1}}(1+\epsilon) \\ (1+\epsilon) + \ldots + u_{n-k,n-k+1}\widehat{x_{n-k+1}}(1+\epsilon)^{k-1}) \end{array}}{u_{n-k,n-k}(1+\epsilon)^k}$$

Stability of LU

$$(1+\epsilon)^k \approx (1+k\epsilon)$$

$$\delta U \approx \begin{bmatrix} n & 1 & 2 & \dots & n-1 \\ 0 & n-1 & 1 & \dots & n-2 \\ 0 & 0 & n-2 & \dots & n-2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \epsilon$$

Stability of LU

$$(1+\epsilon)^k \approx (1+k\epsilon)$$

$$\delta U \approx \begin{bmatrix} n & 1 & 2 & \dots & n-1 \\ 0 & n-1 & 1 & \dots & n-2 \\ 0 & 0 & n-2 & \dots & n-2 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \epsilon$$

$\exists \delta U$ with $\|\delta U\| \leq M\epsilon$

$$(U + \delta U)\hat{x} = b$$

Backsubstitution algorithm is stable

Stability of LU

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\kappa(A) = 2.618\ldots, \ \epsilon = 10^{-16}$$

$$\delta A = \begin{bmatrix} 10^{-20} & 0 \\ 0 & 0 \end{bmatrix}$$

Stability of LU

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\kappa(A) = 2.618\ldots, \ \epsilon = 10^{-16}$$

$$\delta A = \begin{bmatrix} 10^{-20} & 0 \\ 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}, \ U = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1 - 10^{-20} \end{bmatrix}$$

Stability of LU

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\kappa(A) = 2.618\ldots, \ \epsilon = 10^{-16}$$

$$\delta A = \begin{bmatrix} 10^{-20} & 0 \\ 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}, \ U = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1 - 10^{-20} \end{bmatrix}$$

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}, \ \hat{U} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & -10^{-20} \end{bmatrix}, \ \hat{L}\hat{U} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 0 \end{bmatrix}$$

LU factorization is not stable

Pivoting

Partial pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$, $a_{i,k} = \max\{|a_{j,k}|, j = k..n\}$

Pivoting

Partial pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$, $a_{i,k} = \max\{|a_{j,k}|, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A = U$$

$$PA = LU$$

Pivoting

Partial pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$, $a_{i,k} = \max\{|a_{j,k}|, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A = U$$

$$PA = LU$$

Complete pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$
$a_{1:n,k} \leftrightarrow a_{1:n,j}$, $a_{i,j} = \max\{|a_{j,k}|, i = k..n, j = k..n\}$

Pivoting

Partial pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$, $a_{i,k} = \max\{|a_{j,k}|, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A = U$$

$$PA = LU$$

Complete pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$
$a_{1:n,k} \leftrightarrow a_{1:n,j}$, $a_{i,j} = \max\{|a_{j,k}|, i = k..n, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A P_1^{'} P_2^{'} \ldots P_n^{'} = U$$

$$PAP^{'} = LU$$

Pivoting

Partial pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$, $a_{i,k} = \max\{|a_{j,k}|, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A = U$$

$$PA = LU$$

Complete pivoting: $a_{k,1:n} \leftrightarrow a_{i,1:n}$
$a_{1:n,k} \leftrightarrow a_{1:n,j}$, $a_{i,j} = \max\{|a_{j,k}|, i = k..n, j = k..n\}$

$$\widehat{L_n} P_n \widehat{L_{n-1}} P_{n-1} \ldots \widehat{L_1} P_1 A P_1' P_2' \ldots P_n' = U$$

$$PAP' = LU$$

Operations count: Partial pivoting $= \mathcal{O}(n^2)$, Complete pivoting$= \mathcal{O}(n^3)$

Existence and uniqueness

Existence:

$$PA = LU$$

If $\nexists a_{i,k} \neq 0$ for $i = k..n \Rightarrow \det A = 0$

Existence and uniqueness

Existence:

$$PA = LU$$

If $\nexists a_{i,k} \neq 0$ for $i = k..n \Rightarrow \det A = 0$

$$\det A \neq 0 \Rightarrow \forall A,\ \exists P,\ L,\ U \text{ so that } PA = LU$$

Existence and uniqueness

Existence:

$$PA = LU$$

If $\nexists a_{i,k} \neq 0$ for $i = k..n \Rightarrow \det A = 0$

$$\det A \neq 0 \Rightarrow \forall A,\ \exists P,\ L,\ U \text{ so that } PA = LU$$

Uniqueness:

$$A = LU = \hat{L}\hat{U}$$

Existence and uniqueness

Existence:

$$PA = LU$$

If $\nexists a_{i,k} \neq 0$ for $i = k..n \Rightarrow \det A = 0$

$$\det A \neq 0 \Rightarrow \forall A,\ \exists P,\ L,\ U \text{ so that } PA = LU$$

Uniqueness:

$$A = LU = \hat{L}\hat{U}$$

$$\hat{L}^{-1}L = \begin{bmatrix} 1 & 0 & \dots & 0 \\ x & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ x & x & \dots & 1 \end{bmatrix} = \begin{bmatrix} x & x & \dots & x \\ 0 & x & \dots & x \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x \end{bmatrix} = \hat{U}U^{-1}$$

$$L = \hat{L},\ U = \hat{U}$$

Stability of LU with pivoting

$$A = LU$$

$$\hat{L}\hat{U} = A + \delta A$$

$$\frac{\|\delta A\|}{\|L\|\|U\|} = \mathcal{O}(\epsilon)$$

Stability of LU with pivoting

$$A = LU$$

$$\hat{L}\hat{U} = A + \delta A$$

$$\frac{\|\delta A\|}{\|L\|\|U\|} = \mathcal{O}(\epsilon)$$

$$|u_{i,j}| \leq 1 \Rightarrow \|U\| = \mathcal{O}(1)$$

$$\rho = \max_{i,j,a_{i,j} \neq 0} \frac{u_{i,j}}{a_{i,j}}$$

Stability of LU with pivoting

$$A = LU$$

$$\hat{L}\hat{U} = A + \delta A$$

$$\frac{\|\delta A\|}{\|L\|\|U\|} = \mathcal{O}(\epsilon)$$

$$|u_{i,j}| \leq 1 \Rightarrow \|U\| = \mathcal{O}(1)$$

$$\rho = \max_{i,j,a_{i,j} \neq 0} \frac{u_{i,j}}{a_{i,j}}$$

$$\|U\| = \mathcal{O}(\rho)\|A\|$$

$$\frac{\|\delta A\|}{\|A\|} = \mathcal{O}(\rho\epsilon)$$

Stability of LU with pivoting

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

Stability of LU with pivoting

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

Stability of LU with pivoting

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

$$\rho = 2^{n-1} = \max_A \rho(A)$$

Stability of LU with pivoting

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

$$\rho = 2^{n-1} = \max_{A} \rho(A)$$

$$P(\rho(A) \geq \sqrt{n}) = \mathcal{O}(e^{-n}) \; ?$$

Crout factorization:

$$A\widehat{U_1}\dots\widehat{U_n} = L$$

$$A = LU$$

$$L = \begin{bmatrix} x & 0 & 0 & \dots & 0 \\ x & x & 0 & \dots & 0 \\ x & x & x & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & x \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & x & x & \dots & x \\ 0 & 1 & x & \dots & x \\ 0 & 0 & 1 & \dots & x \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$Ax = b \Rightarrow LUx = b \Rightarrow Ux = y$$

LDU factorization

Apply both Doolittle and Crout:

$$\widehat{L_n} \dots \widehat{L_1} A \widehat{U_1} \dots \widehat{U_n} = D$$

LDU factorization

Apply both Doolittle and Crout:

$$\widehat{L_n} \ldots \widehat{L_1} A \widehat{U_1} \ldots \widehat{U_n} = D$$

$$A = LDU$$

$$L = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ x & 1 & 0 & \ldots & 0 \\ x & x & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x & x & x & \ldots & 1 \end{bmatrix} \quad U = \begin{bmatrix} 1 & x & x & \ldots & x \\ 0 & 1 & x & \ldots & x \\ 0 & 0 & 1 & \ldots & x \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

$$D = \operatorname{diag}\left(a_{i,i}\right)$$

LDU factorization

Apply both Doolittle and Crout:

$$\widehat{L_n}\dots\widehat{L_1}A\widehat{U_1}\dots\widehat{U_n} = D$$

$$A = LDU$$

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ x & 1 & 0 & \dots & 0 \\ x & x & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & 1 \end{bmatrix} U = \begin{bmatrix} 1 & x & x & \dots & x \\ 0 & 1 & x & \dots & x \\ 0 & 0 & 1 & \dots & x \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$D = \mathrm{diag}\left(a_{i,i}\right)$$

Doolittle: $A = (LD)U$
Crout: $A = L(DU)$

Cholesky factorization

Hermitian matrix: $A = A^*$, $\Rightarrow a_{i,i} \in \mathbb{R}$

$$x^* A y = \overline{y^* A^* x}$$

Cholesky factorization

Hermitian matrix: $A = A^*$, $\Rightarrow a_{i,i} \in \mathbb{R}$

$$x^* A y = \overline{y^* A^* x}$$

Positive definite:

$$x^* A x > 0, \ \forall \ x \neq 0$$

Cholesky factorization

Hermitian matrix: $A = A^*, \Rightarrow a_{i,i} \in \mathbb{R}$

$$x^* A y = \overline{y^* A^* x}$$

Positive definite:

$$x^* A x > 0, \ \forall \ x \neq 0$$

Let $X \in M^{nxm}(\mathbb{K})$ with full rank so $x \neq 0 \Rightarrow Xx \neq 0$

$$(X^* A X)^* = X^* A X$$

$$x^* (X^* A X) x = (Xx)^* A (Xx) > 0, \ Xx \neq 0$$

Cholesky factorization

Hermitian matrix: $A = A^*, \Rightarrow a_{i,i} \in \mathbb{R}$

$$x^* A y = \overline{y^* A^* x}$$

Positive definite:

$$x^* A x > 0, \ \forall \ x \neq 0$$

Let $X \in M^{nxm}(\mathbb{K})$ with full rank so $x \neq 0 \Rightarrow Xx \neq 0$

$$(X^* A X)^* = X^* A X$$

$$x^*(X^* A X)x = (Xx)^* A (Xx) > 0, \ Xx \neq 0$$

$X^* A X$ is Hermitian and positive definite In particular, the main sub-matrices of a are positive definite, $a_{i,i} > 0$

Cholesky factorization

$$A = \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix}$$

Cholesky factorization

$$A = \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ -\frac{w}{a} & I \end{bmatrix} \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} a & w^* \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

Cholesky factorization

$$A = \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ -\frac{w}{a} & I \end{bmatrix} \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} a & w^* \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$\begin{bmatrix} a & w^* \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} 1 & -\frac{w^*}{a} \\ 0 & I \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

Cholesky factorization

$$A = \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ -\frac{w}{a} & I \end{bmatrix} \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} a & w^* \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$\begin{bmatrix} a & w^* \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} 1 & -\frac{w^*}{a} \\ 0 & I \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{\sqrt{a}} & 0 \\ -\frac{w}{\sqrt{a}} & \sqrt{a}I \end{bmatrix} \begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{a}} & -\frac{w^*}{\sqrt{a}} \\ 0 & \sqrt{a}I \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & aA_2 - ww^* \end{bmatrix}$$

Positive definite matrices

Let $\alpha \in \mathbb{R}_+^*$, A positive definite

Positive definite matrices

Let $\alpha \in \mathbb{R}_+^*$, A positive definite

$$(\alpha A)^* = \alpha A^* = \alpha A,\ x \neq 0 \Rightarrow x^*(\alpha A)x = \alpha(x^* A x) > 0$$

αA is positive definite

Positive definite matrices

Let $\alpha \in \mathbb{R}_+^*$, A positive definite

$$(\alpha A)^* = \alpha A^* = \alpha A, \ x \neq 0 \Rightarrow x^*(\alpha A)x = \alpha(x^* A x) > 0$$

αA is positive definite

$$R^* A R = \begin{bmatrix} 1 & 0 \\ 0 & a A_2 - w w^* \end{bmatrix}$$

Positive definite matrices

Let $\alpha \in \mathbb{R}_+^*$, A positive definite

$$(\alpha A)^* = \alpha A^* = \alpha A,\ x \neq 0 \Rightarrow x^*(\alpha A)x = \alpha(x^* A x) > 0$$

αA is positive definite

$$R^* A R = \begin{bmatrix} 1 & 0 \\ 0 & aA_2 - ww^* \end{bmatrix}$$

$aA_2 - ww^*$ is positive definite
$A_2 - \frac{ww^*}{a}$ is positive definite

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = L_1 A^{(1)} L_1^*$$

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = L_1 A^{(1)} L_1^*$$

$A^{(1)}$ is positive definite

$$A = L_1 L_2 A^{(2)} L_2^* L_1^*$$

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = L_1 A^{(1)} L_1^*$$

$A^{(1)}$ is positive definite

$$A = L_1 L_2 A^{(2)} L_2^* L_1^*$$

$$A = L_1 L_2 \dots L_n A^{(n)} L_n^* \dots L_2^* L_1^*$$

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = L_1 A^{(1)} L_1^*$$

$A^{(1)}$ is positive definite

$$A = L_1 L_2 A^{(2)} L_2^* L_1^*$$

$$A = L_1 L_2 \dots L_n A^{(n)} L_n^* \dots L_2^* L_1^*$$

$$A^{(n)} = I, \; L_1 L_2 \dots L_n = L \Rightarrow L^* = L_n^* \dots L_2^* L_1^*$$

Cholesky factorization

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{w}{\sqrt{a}} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{w^*}{\sqrt{a}} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = L_1 A^{(1)} L_1^*$$

$A^{(1)}$ is positive definite

$$A = L_1 L_2 A^{(2)} L_2^* L_1^*$$

$$A = L_1 L_2 \ldots L_n A^{(n)} L_n^* \ldots L_2^* L_1^*$$

$$A^{(n)} = I, \; L_1 L_2 \ldots L_n = L \Rightarrow L^* = L_n^* \ldots L_2^* L_1^*$$

$$A = LL^*$$

Cholesky factorization

$$L = \begin{bmatrix} \sqrt{a_{1,1}} & 0 & 0 & \dots & 0 \\ x & \sqrt{a_{2,2}} & 0 & \dots & 0 \\ x & x & \sqrt{a_{3,3}} & \dots & 0 \\ \hdotsfor{5} \\ x & x & x & \dots & \sqrt{a_{n,n}} \end{bmatrix}$$

Cholesky factorization

$$L = \begin{bmatrix} \sqrt{a_{1,1}} & 0 & 0 & \dots & 0 \\ x & \sqrt{a_{2,2}} & 0 & \dots & 0 \\ x & x & \sqrt{a_{3,3}} & \dots & 0 \\ \hdotsfor{5} \\ x & x & x & \dots & \sqrt{a_{n,n}} \end{bmatrix}$$

$\forall A$ positive definite $\exists! L$ so $A = LL^*$

Cholesky factorization

$$L = \begin{bmatrix} \sqrt{a_{1,1}} & 0 & 0 & \dots & 0 \\ x & \sqrt{a_{2,2}} & 0 & \dots & 0 \\ x & x & \sqrt{a_{3,3}} & \dots & 0 \\ \hdotsfor{5} \\ x & x & x & \dots & \sqrt{a_{n,n}} \end{bmatrix}$$

$\forall A$ positive definite $\exists! L$ so $A = LL^*$
Operations count:

$$\sum_{k=1}^{n} \sum_{j=k+1}^{n} 2(n-j) \approx \frac{1}{3}n^3 = \mathcal{O}(n^3)$$

Cholesky factorization

$$L = \begin{bmatrix} \sqrt{a_{1,1}} & 0 & 0 & \dots & 0 \\ x & \sqrt{a_{2,2}} & 0 & \dots & 0 \\ x & x & \sqrt{a_{3,3}} & \dots & 0 \\ \hdotsfor{5} \\ x & x & x & \dots & \sqrt{a_{n,n}} \end{bmatrix}$$

$\forall A$ positive definite $\exists! L$ so $A = LL^*$
Operations count:

$$\sum_{k=1}^{n} \sum_{j=k+1}^{n} 2(n-j) \approx \frac{1}{3}n^3 = \mathcal{O}(n^3)$$

There are n square roots

Diagonal Cholesky

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} & 0 \\ \frac{w}{a} & I \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} a & \frac{w^*}{a} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

Diagonal Cholesky

$$\begin{bmatrix} a & w^* \\ w & A_2 \end{bmatrix} = \begin{bmatrix} & 0 \\ \frac{w}{a} & I \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix} \begin{bmatrix} a & \frac{w^*}{a} \\ 0 & A_2 - \frac{ww^*}{a} \end{bmatrix}$$

$$A = \widehat{L} D \widehat{L^*} = (L\sqrt{D}) A (L\sqrt{D})^*$$

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ x & 1 & 0 & \dots & 0 \\ x & x & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} a_{1,1} & 0 & 0 & \dots & 0 \\ 0 & a_{2,2} & 0 & \dots & 0 \\ 0 & 0 & a_{3,3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{bmatrix}$$

Stability of Cholesky

Let $\| \cdot \|_F$ be the Frobenius norm

Stability of Cholesky

Let $\| \cdot \|_F$ be the Frobenius norm

$$\|A\|_F = \operatorname{tr}(AA^*) = \operatorname{tr}((LL^*)(LL^*)^*) = \operatorname{tr}(LL^*)\operatorname{tr}(LL^*) = \|L\|_F \|L^*\|_F$$

Stability of Cholesky

Let $\|\cdot\|_F$ be the Frobenius norm

$$\|A\|_F = \operatorname{tr}(AA^*) = \operatorname{tr}((LL^*)(LL^*)^*) = \operatorname{tr}(LL^*)\operatorname{tr}(LL^*) = \|L\|_F\|L^*\|_F$$

$$\frac{\|\delta A\|_F}{\|L\|_F\|L^*\|_F} = \frac{\|\delta A\|_F}{\|A\|_F} = \mathcal{O}(\epsilon)$$

Stability of Cholesky

Let $\|\cdot\|_F$ be the Frobenius norm

$\|A\|_F = \operatorname{tr}(AA^*) = \operatorname{tr}((LL^*)(LL^*)^*) = \operatorname{tr}(LL^*)\operatorname{tr}(LL^*) = \|L\|_F\|L^*\|_F$

$$\frac{\|\delta A\|_F}{\|L\|_F\|L^*\|_F} = \frac{\|\delta A\|_F}{\|A\|_F} = \mathcal{O}(\epsilon)$$

Cholesky factorization is stable

Gram-Schmidt orthogonalization

Let $(v_1,\ v_2,\ \dots\ v_k)$ be linearly independent
$V = \text{span}\,(v_1,\ v_2,\ \dots\ v_n)$

Gram-Schmidt orthogonalization

Let $(v_1,\ v_2,\ \dots\ v_k)$ be linearly independent
$V = \operatorname{span}(v_1,\ v_2,\ \dots\ v_n)$

$$u_1 = v_1$$

$$u_2 = v_2 - \frac{u_1^* v_2}{u_1^* u_1} u_1$$

$$u_3 = v_3 - \frac{u_1^* v_3}{u_1^* u_1} u_1 - \frac{u_2^* v_3}{u_2^* u_2} u_2$$

$$\dots$$

$$u_n = v_n - \frac{u_1^* v_n}{u_1^* u_1} u_1 - \dots - \frac{u_{n-1}^* v_3}{u_{n-1}^* u_2} u_{n-1}$$

Gram-Schmidt orthogonalization

$$v_1 = u_1$$

$$v_2 = u_2 + \frac{u_1^* v_2}{u_1^* u_1} u_1$$

$$v_3 = u_3 + \frac{u_1^* v_3}{u_1^* u_1} u_1 + \frac{u_2^* v_3}{u_2^* u_2} u_2$$

$$\ldots$$

$$v_n = u_n + \frac{u_1^* v_n}{u_1^* u_1} u_1 + \ldots + \frac{u_{n-1}^* v_3}{u_{n-1}^* u_2} u_{n-1}$$

Gram-Schmidt orthogonalization

$$v_1 = u_1$$
$$v_2 = u_2 + \frac{u_1^* v_2}{u_1^* u_1} u_1$$
$$v_3 = u_3 + \frac{u_1^* v_3}{u_1^* u_1} u_1 + \frac{u_2^* v_3}{u_2^* u_2} u_2$$
$$\dots$$
$$v_n = u_n + \frac{u_1^* v_n}{u_1^* u_1} u_1 + \ldots + \frac{u_{n-1}^* v_3}{u_{n-1}^* u_2} u_{n-1}$$

$$V = \mathrm{span}\,(u_1,\ u_2,\ \ldots\ u_n),\ u_i \neq 0$$

Orthogonality of the Gram-Schmidt

$$u_1^* u_2 = u_1^* v_2 - \frac{u_1^* v_2}{u_1^* u_1} u_1^* u_1 = 0$$

$$u_1^* u_3 = u_1^* v_3 - \frac{u_1^* v_3}{u_1^* u_1} u_1^* u_1 - \frac{u_2^* v_3}{u_2^* u_2} u_1^* u_2 = 0$$

$$u_2^* u_3 = u_2^* v_3 - \frac{u_1^* v_3}{u_1^* u_1} u_2^* u_1 - \frac{u_2^* v_3}{u_2^* u_2} u_2^* u_2 = 0$$

$$\cdots$$

$$u_i^* u_j = u_i^* v_j - \frac{u_1^* v_j}{u_1^* u_1} u_i^* u_1 - \ldots - \frac{u_{j-1}^* v_j}{u_{j-1}^* u_{j-1}} u_i^* u_{j-1} = 0$$

Orthogonality of the Gram-Schmidt

$$u_1^* u_2 = u_1^* v_2 - \frac{u_1^* v_2}{u_1^* u_1} u_1^* u_1 = 0$$

$$u_1^* u_3 = u_1^* v_3 - \frac{u_1^* v_3}{u_1^* u_1} u_1^* u_1 - \frac{u_2^* v_3}{u_2^* u_2} u_1^* u_2 = 0$$

$$u_2^* u_3 = u_2^* v_3 - \frac{u_1^* v_3}{u_1^* u_1} u_2^* u_1 - \frac{u_2^* v_3}{u_2^* u_2} u_2^* u_2 = 0$$

$$\cdots$$

$$u_i^* u_j = u_i^* v_j - \frac{u_1^* v_j}{u_1^* u_1} u_i^* u_1 - \ldots - \frac{u_{j-1}^* v_j}{u_{j-1}^* u_{j-1}} u_i^* u_{j-1} = 0$$

$$u_i^* u_j = 0 \; \forall \; i \neq j$$

QR Factorization

Let $A = \begin{bmatrix} v_1, & v_2, \ldots, v_k \end{bmatrix}$ with full rank

QR Factorization

Let $A = \begin{bmatrix} v_1, & v_2, \dots, v_k \end{bmatrix}$ with full rank

$$v_1 = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$v_i = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} \frac{u_1^* v_i}{u_1^* u_1} \\ \frac{u_2^* v_i}{u_2^* u_2} \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

QR Factorization

$$A = \begin{bmatrix} \frac{u_1}{\|u_1\|} & \frac{u_2}{\|u_2\|} & \cdots & \frac{u_n}{\|u_n\|} \end{bmatrix} \begin{bmatrix} \|u_1\| & u_1^* v_2 & u_1^* v_3 & \dots & u_1^* v_k \\ 0 & \|u_2\| & u_2^* v_3 & \dots & u_2^* v_k \\ 0 & 0 & \|u_3\| & \dots & u_3^* v_k \\ \hdotsfor{5} \\ 0 & 0 & 0 & \dots & \|u_n\| \end{bmatrix}$$

QR Factorization

$$A = \begin{bmatrix} \frac{u_1}{\|u_1\|} & \frac{u_2}{\|u_2\|} & \cdots & \frac{u_n}{\|u_n\|} \end{bmatrix} \begin{bmatrix} \|u_1\| & u_1^* v_2 & u_1^* v_3 & \dots & u_1^* v_k \\ 0 & \|u_2\| & u_2^* v_3 & \dots & u_2^* v_k \\ 0 & 0 & \|u_3\| & \dots & u_3^* v_k \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \|u_n\| \end{bmatrix}$$

Let $q_i = \frac{u_i}{\|u_i\|}$, $\widehat{Q} = \left[q_1,\ q_2, \dots, q_k\right]$

$$A = \widehat{Q}\widehat{R}$$

$\widehat{Q}$ is orthogonal and $\widehat{R}$ is upper-triangular

Full QR factorization

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
Let $m \geq n$

Full QR factorization

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
Let $m \geq n$
Linear algebra and Gram-Schmidt:

$$Q = \left[\widehat{Q}, q_{n+1}, q_{n+2}, \ldots, q_m\right]$$

$Q \in M^{mxm}(\mathbb{K})$ orthogonal

Full QR factorization

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
Let $m \geq n$
Linear algebra and Gram-Schmidt:

$$Q = \left[\widehat{Q}, q_{n+1}, q_{n+2}, \ldots, q_m\right]$$

$Q \in M^{mxm}(\mathbb{K})$ orthogonal

$$R = \begin{bmatrix} \widehat{R} \\ 0 \end{bmatrix}$$

$Q \in M^{mxn}(\mathbb{K})$ upper-triangular

Full QR factorization

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
Let $m \geq n$
Linear algebra and Gram-Schmidt:

$$Q = \left[\widehat{Q}, q_{n+1}, q_{n+2}, \ldots, q_m\right]$$

$Q \in M^{mxm}(\mathbb{K})$ orthogonal

$$R = \begin{bmatrix} \widehat{R} \\ 0 \end{bmatrix}$$

$Q \in M^{mxn}(\mathbb{K})$ upper-triangular

$$A = QR$$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with rank n-1

$$v_n = \sum_{i=1}^{n-1} \alpha_i v_i$$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with rank n-1

$$v_n = \sum_{i=1}^{n-1} \alpha_i v_i$$

$$\text{span}\left(u_1 \ldots u_{n-1}\right) = \text{span}\left(v_1 \ldots v_{n-1}\right) = \text{span}\left(v_1 \ldots v_n\right)$$

$$\text{span}\left(u_1 \ldots u_{n-1}\right) = \text{span}\left(u_1 \ldots u_n\right)$$

$$u_n \in \text{span}\left(u_1 \ldots u_{n-1}\right)$$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with rank n-1

$$v_n = \sum_{i=1}^{n-1} \alpha_i v_i$$

$$\text{span}\left(u_1 \dots u_{n-1}\right) = \text{span}\left(v_1 \dots v_{n-1}\right) = \text{span}\left(v_1 \dots v_n\right)$$

$$\text{span}\left(u_1 \dots u_{n-1}\right) = \text{span}\left(u_1 \dots u_n\right)$$

$$u_n \in \text{span}\left(u_1 \dots u_{n-1}\right)$$

$$u_i^* u_n = 0 \; \forall \; i \Rightarrow u_n = 0$$

Existence and uniqueness of QR

$$\widehat{Q} = \begin{bmatrix} q_1 & \dots & q_{n-1} & 0 \end{bmatrix}$$

$$\widehat{R} = \begin{bmatrix} x & x & x & \dots & \beta_1 \\ 0 & x & x & \dots & \beta_2 \\ 0 & 0 & x & \dots & \beta_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Existence and uniqueness of QR

$$\widehat{Q} = \begin{bmatrix} q_1 & \dots & q_{n-1} & 0 \end{bmatrix}$$

$$\widehat{R} = \begin{bmatrix} x & x & x & \dots & \beta_1 \\ 0 & x & x & \dots & \beta_2 \\ 0 & 0 & x & \dots & \beta_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Induction for rank $\leq n-2$

$$\forall A \in M^{mxn}(\mathbb{K})\ \exists \widehat{Q},\ \widehat{R},\ A = \widehat{Q}\widehat{R}$$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
The Gram-Schmidt process $A = \widehat{Q}\widehat{R}$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
The Gram-Schmidt process $A = \widehat{Q}\widehat{R}$
Induction: $n = 1$, $v_1 = zq_1 \frac{\|u_1\|}{z}$ wirh $|z| = 1$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
The Gram-Schmidt process $A = \widehat{Q}\widehat{R}$
Induction: $n = 1$, $v_1 = zq_1\frac{\|u_1\|}{z}$ wirh $|z| = 1$
For n, $q_1, \ \ldots \ q_{n-1}$ are fixed $\Rightarrow q_n = \frac{zu_n}{\|u_n\|}$ with $|z| = 1$

Existence and uniqueness of QR

Let $A \in M^{mxn}(\mathbb{K})$ with full rank
The Gram-Schmidt process $A = \widehat{Q}\widehat{R}$
Induction: $n = 1$, $v_1 = zq_1 \frac{\|u_1\|}{z}$ wirh $|z| = 1$
For n, $q_1, \ \ldots \ q_{n-1}$ are fixed $\Rightarrow q_n = \frac{zu_n}{\|u_n\|}$ with $|z| = 1$

$$A = \widehat{Q'}\widehat{R'}$$

$$r_{i,i} = zr^{'}_{i,i}, \ |z| = 1$$

Instability of Gram-Schmidt

$$a_1 = [1 \ \epsilon_0 \ 0 \ 0]^*, \ \ a_2 = [1 \ 0 \ \epsilon_0 \ 0]^*, \ \ a_3 = [1 \ 0 \ 0 \ \epsilon_0]^*$$

Instability of Gram-Schmidt

$$a_1 = [1 \ \epsilon_0 \ 0 \ 0]^*, \ a_2 = [1 \ 0 \ \epsilon_0 \ 0]^*, \ a_3 = [1 \ 0 \ 0 \ \epsilon_0]^*$$

$$\epsilon_0 \epsilon, \ 1 + \epsilon_0^2 \approx 1$$

Instability of Gram-Schmidt

$$a_1 = [1\ \epsilon_0\ 0\ 0]^*,\ a_2 = [1\ 0\ \epsilon_0\ 0]^*,\ a_3 = [1\ 0\ 0\ \epsilon_0]^*$$

$$\epsilon_0 \epsilon,\ 1 + \epsilon_0^2 \approx 1$$

$$u_1 = a_1 = [1\ \epsilon_0\ 0\ 0]^*$$

$$u_2 = a_2 - \frac{1}{1+\epsilon_0^2} u_1 = [0\ \ -\epsilon_0\ \epsilon_0\ 0]^*$$

$$u_3 = a_3 - \frac{1}{1+\epsilon_0^2} u_1 - \frac{0}{\epsilon_0\sqrt{2}} u_2 = [0\ \ -\epsilon_0\ 0\ \epsilon_0]^*$$

Instability of Gram-Schmidt

$$q_1 = [1 \; \epsilon_0 \; 0 \; 0]^*$$

$$q_2 = \left[0 \; -\frac{1}{\sqrt{2}} \; \frac{1}{\sqrt{2}} \; 0\right]^*$$

$$q_3 = \left[0 \; -\frac{1}{\sqrt{2}} \; 0 \; \frac{1}{\sqrt{2}}\right]^*$$

Instability of Gram-Schmidt

$$q_1 = [1 \ \epsilon_0 \ 0 \ 0]^*$$

$$q_2 = \left[0 \ -\frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} \ 0\right]^*$$

$$q_3 = \left[0 \ -\frac{1}{\sqrt{2}} \ 0 \ \frac{1}{\sqrt{2}}\right]^*$$

$$q_2^* q_3 = \frac{1}{2} \neq 0$$

Instability of Gram-Schmidt

$$q_1 = [1 \;\; \epsilon_0 \;\; 0 \;\; 0]^*$$

$$q_2 = \left[0 \;\; -\tfrac{1}{\sqrt{2}} \;\; \tfrac{1}{\sqrt{2}} \;\; 0\right]^*$$

$$q_3 = \left[0 \;\; -\tfrac{1}{\sqrt{2}} \;\; 0 \;\; \tfrac{1}{\sqrt{2}}\right]^*$$

$$q_2^* q_3 = \tfrac{1}{2} \neq 0$$

$$\widehat{Q'} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon_0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Instability of Gram-Schmidt

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \delta A = \begin{bmatrix} 0 & 0 & 0 \\ \epsilon_0 & 0 & 0 \\ 0 & \epsilon_0 & 0 \\ 0 & 0 & \epsilon_0 \end{bmatrix}$$

Instability of Gram-Schmidt

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \delta A = \begin{bmatrix} 0 & 0 & 0 \\ \epsilon_0 & 0 & 0 \\ 0 & \epsilon_0 & 0 \\ 0 & 0 & \epsilon_0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \widehat{Q}\widehat{R}$$

Instability of Gram-Schmidt

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \delta A = \begin{bmatrix} 0 & 0 & 0 \\ \epsilon_0 & 0 & 0 \\ 0 & \epsilon_0 & 0 \\ 0 & 0 & \epsilon_0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \widehat{Q}\widehat{R}$$

$$\|\delta A\| = \mathcal{O}(\epsilon)$$

$$\|\widehat{Q} - \widehat{Q'}\| = \mathcal{O}\left(\frac{1}{\sqrt{2}}\right)$$

Modified Gram-Schmidt

$$u_k^{(1)} = v_k - \frac{u_1^* v_k}{u_1^* u_1} u_1$$

$$u_k^{(2)} = u_k^{(2)} - \frac{u_1^* u_k^{(2)}}{u_1^* u_1} u_1$$

$$u_k^{(3)} = u_k^{(3)} - \frac{u_1^* u_k^{(2)}}{u_1^* u_1} u_1$$

$$\cdots$$

$$u_k^{(k-1)} = u_k^{(k-1)} - \frac{u_1^* u_k^{(k-2)}}{u_1^* u_1} u_1$$

Modified Gram-Schmidt

$$u_k^{(1)} = v_k - \frac{u_1^* v_k}{u_1^* u_1} u_1$$

$$u_k^{(2)} = u_k^{(2)} - \frac{u_1^* u_k^{(2)}}{u_1^* u_1} u_1$$

$$u_k^{(3)} = u_k^{(3)} - \frac{u_1^* u_k^{(2)}}{u_1^* u_1} u_1$$

$$\cdots$$

$$u_k^{(k-1)} = u_k^{(k-1)} - \frac{u_1^* u_k^{(k-2)}}{u_1^* u_1} u_1$$

$$u_k = u_k^{(k-1)}$$

Modified Gram-Schmidt

$$u_1 = a_1 = [1 \ \epsilon_0 \ 0 \ 0]^*$$

$$u_2 = a_2 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \ -\epsilon_0 \ \epsilon_0 \ 0]^*$$

$$u_3^{(1)} = a_3 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \ -\epsilon_0 \ 0 \ \epsilon_0]^*$$

$$u_3^{(2)} = u_3^{(1)} - \frac{\epsilon_0^2}{2\epsilon_0^2} u_2 = \left[0 \ -\frac{\epsilon_0}{2} \ -\frac{\epsilon_0}{2} \ \epsilon_0\right]^*$$

Modified Gram-Schmidt

$$u_1 = a_1 = [1 \ \epsilon_0 \ 0 \ 0]^*$$

$$u_2 = a_2 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \ -\epsilon_0 \ \epsilon_0 \ 0]^*$$

$$u_3^{(1)} = a_3 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \ -\epsilon_0 \ 0 \ \epsilon_0]^*$$

$$u_3^{(2)} = u_3^{(1)} - \frac{\epsilon_0^2}{2\epsilon_0^2} u_2 = \left[0 \ -\frac{\epsilon_0}{2} \ -\frac{\epsilon_0}{2} \ \epsilon_0\right]^*$$

$$q_1 = [1 \ \epsilon_0 \ 0 \ 0], \ q_2 = \frac{1}{\sqrt{2}}[0 \ -1 \ 1 \ 0], q_3 = \frac{1}{\sqrt{6}}[0 \ -1 \ -1 \ 2]$$

Modified Gram-Schmidt

$$u_1 = a_1 = [1 \;\; \epsilon_0 \;\; 0 \;\; 0]^*$$

$$u_2 = a_2 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \;\; -\epsilon_0 \;\; \epsilon_0 \;\; 0]^*$$

$$u_3^{(1)} = a_3 - \frac{1}{1+\epsilon_0^2} u_1 = [0 \;\; -\epsilon_0 \;\; 0 \;\; \epsilon_0]^*$$

$$u_3^{(2)} = u_3^{(1)} - \frac{\epsilon_0^2}{2\epsilon_0^2} u_2 = \left[0 \;\; -\frac{\epsilon_0}{2} \;\; -\frac{\epsilon_0}{2} \;\; \epsilon_0\right]^*$$

$$q_1 = [1 \;\; \epsilon_0 \;\; 0 \;\; 0], \;\; q_2 = \frac{1}{\sqrt{2}}[0 \;\; -1 \;\; 1 \;\; 0], q_3 = \frac{1}{\sqrt{6}}[0 \;\; -1 \;\; -1 \;\; 2]$$

$$q_1^* q_2 = \frac{\epsilon_0}{\sqrt{2}}, \;\; q_1^* q_3 = \frac{\epsilon_0}{\sqrt{6}}, \;\; q_2^* q_3 = 0$$

Householder's algorithm

Let $A \in M^{mxn}(\mathbb{K})$ with full rank

Householder's algorithm

Let $A \in M^{mxn}(\mathbb{K})$ with full rank

$$Q_1 A = A_1$$
$$Q_2 A_1 = A_2$$
$$\ldots$$
$$Q_n A_{n-1} = A_n = R$$

Householder's algorithm

Let $A \in M^{mxn}(\mathbb{K})$ with full rank

$$Q_1 A = A_1$$
$$Q_2 A_1 = A_2$$
$$\dots$$
$$Q_n A_{n-1} = A_n = R$$

$$Q_n Q_{n-1} \dots Q_1 A = R$$

Householder's algorithm

Let $A \in M^{mxn}(\mathbb{K})$ with full rank

$$Q_1 A = A_1$$
$$Q_2 A_1 = A_2$$
$$\dots$$
$$Q_n A_{n-1} = A_n = R$$

$$Q_n Q_{n-1} \dots Q_1 A = R$$

$$Q_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & F \end{bmatrix}, \; FF^* = I$$

Householder's algorithm

$$Fx = \|x\| e_1$$

Householder's algorithm

$$Fx = \|x\| e_1$$

Consider $v = x - \|x\| e_1$ and $H \perp v$

Householder's algorithm

$$Fx = \|x\| e_1$$

Consider $v = x - \|x\| e_1$ and $H \perp v$
The projection of x on H:

$$x - \frac{v^* x}{v^* v} v$$

Householder's algorithm

$$Fx = \|x\| e_1$$

Consider $v = x - \|x\| e_1$ and $H \perp v$
The projection of x on H:

$$x - \frac{v^* x}{v^* v} v$$

The reflection across H:

$$x - 2\frac{v^* x}{v^* v} v = Ix - 2Iv\frac{v^* x}{v * v} = (I - 2\frac{vv^*}{v^* v})x$$

$$F = I - 2\frac{vv^*}{v^* v}$$

Householder's algorithm

$$A_{k-1} = \begin{bmatrix} x & x & x & \dots & x_1 & \dots & x \\ 0 & x & x & \dots & x_2 & \dots & x \\ 0 & 0 & x & \dots & x_3 & \dots & x \\ \hline 0 & 0 & 0 & \dots & x_k & \dots & x \\ 0 & 0 & 0 & \dots & x_{k+1} & \dots & x \\ \hline 0 & 0 & 0 & \dots & x_m & \dots & x \end{bmatrix} = \begin{bmatrix} A_{1,1} & y_0 & A_{1,2} \\ 0 & y & A_{2,2} \end{bmatrix}$$

Householder's algorithm

$$A_{k-1} = \begin{bmatrix} x & x & x & \dots & x_1 & \dots & x \\ 0 & x & x & \dots & x_2 & \dots & x \\ 0 & 0 & x & \dots & x_3 & \dots & x \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \dots & x_k & \dots & x \\ 0 & 0 & 0 & \dots & x_{k+1} & \dots & x \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \dots & x_m & \dots & x \end{bmatrix} = \begin{bmatrix} A_{1,1} & y_0 & A_{1,2} \\ 0 & y & A_{2,2} \end{bmatrix}$$

$$y = [x_k \; x_{k+1} \; \dots \; x_m]^T$$

Householder's algorithm

$$A_{k-1} = \left[\begin{array}{ccccccc} x & x & x & \dots & x_1 & \dots & x \\ 0 & x & x & \dots & x_2 & \dots & x \\ 0 & 0 & x & \dots & x_3 & \dots & x \\ \hdotsfor{7} \\ 0 & 0 & 0 & \dots & x_k & \dots & x \\ 0 & 0 & 0 & \dots & x_{k+1} & \dots & x \\ \hdotsfor{7} \\ 0 & 0 & 0 & \dots & x_m & \dots & x \end{array}\right] = \begin{bmatrix} A_{1,1} & y_0 & A_{1,2} \\ 0 & y & A_{2,2} \end{bmatrix}$$

$$y = [x_k \ x_{k+1} \ \dots \ x_m]^T$$

$$\begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix} \begin{bmatrix} A_{1,1} & y_0 & A_{1,2} \\ 0 & y & A_{2,2} \end{bmatrix} = \begin{bmatrix} A_{1,1} & y_0 & A_{1,2} \\ 0 & Fy & FA_{2,2} \end{bmatrix}$$

Sign of Householder's reflector

$$Fx = z\|x\|e_1, \ \ |z| = 1$$

Sign of Householder's reflector

$$Fx = z\|x\|e_1,\ \ |z| = 1$$

Numerical stability: $\|x - z\|x\|e_1\| = \|v\| = \max$

Sign of Householder's reflector

$$Fx = z\|x\|e_1, \; |z| = 1$$

Numerical stability: $\|x - z\|x\|e_1\| = \|v\| = \max$

$$x = x_k e_1 + x_0$$

$$Fx = \alpha x$$

Sign of Householder's reflector

$$Fx = z\|x\|e_1, \ |z| = 1$$

Numerical stability: $\|x - z\|x\|e_1\| = \|v\| = \max$

$$x = x_k e_1 + x_0$$

$$Fx = \alpha x$$

$\|x - Fx\|$ is maximum when $\|(x_k - \alpha)e_1 + x_0\| \leq |x_k - \alpha| + \|x_0\|$ is maximum $\Rightarrow |x_k - \alpha| = |x_k| + |\alpha|$

Sign of Householder's reflector

$$Fx = z\|x\|e_1, \;\; |z| = 1$$

Numerical stability: $\|x - z\|x\|e_1\| = \|v\| = \max$

$$x = x_k e_1 + x_0$$

$$Fx = \alpha x$$

$\|x - Fx\|$ is maximum when $\|(x_k - \alpha)e_1 + x_0\| \le |x_k - \alpha| + \|x_0\|$ is maximum $\Rightarrow |x_k - \alpha| = |x_k| + |\alpha|$

$$\operatorname{sgn}(x_k) = \operatorname{sgn}(-\alpha) = \operatorname{sgn}(-z) = \operatorname{sgn}(z) \Rightarrow z = -\operatorname{sgn}(x_k)$$

Computing Q

$$Q_n Q_{n-1} \dots Q_1 A = R$$

Computing Q

$$Q_n Q_{n-1} \dots Q_1 A = R$$

$$A = Q_1^* 1 Q_2^* \dots Q_n^* R$$

$$Q_1^* Q_2^* \dots Q_n^* = Q$$

Computing Q

$$Q_n Q_{n-1} \dots Q_1 A = R$$

$$A = Q_1^* 1 Q_2^* \dots Q_n^* R$$

$$Q_1^* Q_2^* \dots Q_n^* = Q$$

$$I^{(2)} = Q_1 I,\ I^{(3)} = Q_2 I^{(1)},\ \dots\ Q^* = I^{(n)} = Q_n I^{(n-1)} \Rightarrow Q = I^{(n)*}$$

Computing Q

$$Q_n Q_{n-1} \dots Q_1 A = R$$

$$A = Q_1^* 1 Q_2^* \dots Q_n^* R$$

$$Q_1^* Q_2^* \dots Q_n^* = Q$$

$$I^{(2)} = Q_1 I,\ \ I^{(3)} = Q_2 I^{(1)},\ \dots\ \ Q^* = I^{(n)} = Q_n I^{(n-1)} \Rightarrow Q = I^{(n)*}$$

Linear Systems:

$$Ax = b \Rightarrow Q_1 Ax = Q_1 b \Rightarrow \dots Q^* Ax = Rx = Q * b$$

$$Ax = QRx = b \Rightarrow Rx = Q^* b$$

Stability of Householder's algorithm

We know if $QQ^* = I$

$$Qx = y \Rightarrow \widehat{y} = (Q + \delta Q)x, \; \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

Stability of Householder's algorithm

We know if $QQ^* = I$

$$Qx = y \Rightarrow \widehat{y} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$QA = B \Rightarrow \widehat{B} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

Stability of Householder's algorithm

We know if $QQ^* = I$

$$Qx = y \Rightarrow \widehat{y} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$QA = B \Rightarrow \widehat{B} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$Q_n Q_{n-1} \dots Q_1 A = R \Rightarrow \widehat{Q^*} A = \widehat{R}, \ \|\widehat{Q^*} - Q^*\| = \mathcal{O}(\epsilon)$$

Stability of Householder's algorithm

We know if $QQ^* = I$

$$Qx = y \Rightarrow \widehat{y} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$QA = B \Rightarrow \widehat{B} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$Q_n Q_{n-1} \dots Q_1 A = R \Rightarrow \widehat{Q^*} A = \widehat{R}, \ \|\widehat{Q^*} - Q^*\| = \mathcal{O}(\epsilon)$$

Householder's orthogonal triangularization is backward-stable

Stability of Householder's algorithm

We know if $QQ^* = I$

$$Qx = y \Rightarrow \widehat{y} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$QA = B \Rightarrow \widehat{B} = (Q + \delta Q)x, \ \|\delta Q\|_F = \mathcal{O}(\epsilon)$$

$$Q_n Q_{n-1} \dots Q_1 A = R \Rightarrow \widehat{Q^*} A = \widehat{R}, \ \|\widehat{Q^*} - Q^*\| = \mathcal{O}(\epsilon)$$

Householder's orthogonal triangularization is backward-stable

$$A = QR \approx \widehat{Q}\widehat{R} \Rightarrow A + \delta A = Q\widehat{R}, \ \frac{\|\delta A\|}{\|A\|} = \mathcal{O}(n)$$

Given's rotation

$$
\begin{bmatrix}
1 & 0 & \dots & 0 & 0 \\
0 & 1 & \dots & 0 & 0 \\
\hdotsfor{5} \\
0 & 0 & \dots & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_k}{\sqrt{x_{k-1}^2+x_k^2}} \\
0 & 0 & \dots & \frac{-x_k}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}}
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{bmatrix}
=
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \sqrt{x_{k-1}^2 + x_k^2} \\ 0 \end{bmatrix}
$$

Given's rotation

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \hdotsfor{5} \\ 0 & 0 & \dots & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_k}{\sqrt{x_{k-1}^2+x_k^2}} \\ 0 & 0 & \dots & \frac{-x_k}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \sqrt{x_{k-1}^2 + x_k^2} \\ 0 \end{bmatrix}$$

$$G_1 x = G^{(1)} G^{(2)} \dots G^{(k-1)} x = \|x\| e_1$$

Given's rotation

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \hdotsfor{5} \\ 0 & 0 & \dots & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_k}{\sqrt{x_{k-1}^2+x_k^2}} \\ 0 & 0 & \dots & \frac{-x_k}{\sqrt{x_{k-1}^2+x_k^2}} & \frac{x_{k-1}}{\sqrt{x_{k-1}^2+x_k^2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \sqrt{x_{k-1}^2 + x_k^2} \\ 0 \end{bmatrix}$$

$$G_1 x = G^{(1)} G^{(2)} \dots G^{(k-1)} x = \|x\| e_1$$

$$G_n G_{n-1} \dots G_1 A = R$$

Cholesky algorithm for QR

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

Cholesky algorithm for QR

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

$$A^*A = LL^*$$
$$L^{-1}A^*A = L^*$$

Cholesky algorithm for QR

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

$$A^*A = LL^*$$

$$L^{-1}A^*A = L^*$$

$$(L^{-1}A^*)(L^{-1}A^*)^* = L^{-1}A^*AL^{-*} = L^{-1}LL^*L^{-*} = I$$

$$A^* = LL^{-1}A^* \Rightarrow A = (L^{-1}A^*)^*L^*$$

References I

- L. N. Trefethen, David Bau, Numerical linear algebra, SIAM, 1997
- J. Demmel, Applied numerical linear algebra, SIAM, 1997
- P. G. Ciarlet, Introduction to numerical linear algebra and optimization, Cambridge University Press, 1989
- G. W. Stewart, Matrix algorithms volumes I II, SIAM, 1998
- G. W. Stewart, Introduction to matrix computations, Academic Press, 1973
- G. W. Stewart, Ji-guang Sun, Matrix perturbation theory, Academic Press, 1990
- G. H. Golub, C. van Loan, Matrix computations, John Hopkins University Press, 1996

References II

- N. J Higham, Accuracy and stability of numerical algorithms, SIAM, 2002
- J. Wilkinson, The algebraic eigenvalue problem, Oxford University Press, 1988
- C. D. Meyer, Matrix analysis and applied linear algebra, SIAM, 2000