Kinds of Matrices

Symmetric: $A = A^T$

Skew-Symmetric: $A = -A^T$

Hermitian: $A = A^*$ ← conjugate transpose

Unitary (Orthogonal): $Q^*Q = I$, $Q^{-1} = Q^*$ ($Q \in \mathbb{C}^{m\times m}$ unitary, $Q \in \mathbb{R}^{m\times m}$ orthogonal) ; $q_i^* q_j = \delta_{ij}$

Normal Matrix: $A^*A = AA^*$; **Theorem:** a matrix is unitarily diagonalizable iff it is normal $\|Qx\| = \|x\|$

Idempotent: $P^2 = P$

Nilpotent: $P^k = 0$

Notes:

① $Ax = b$:

$$\begin{array}{ccc} b & \xrightarrow{\text{mult. by } A^{-1}} & A^{-1}b \\ \text{coefficients of the} & \xleftarrow[\text{mult. by } A]{} & \text{coefficients of the} \\ \text{expan. of } b \text{ in} & & \text{expan. of } b \text{ in} \\ \{e_1, \dots, e_m\} & & \{a_1, a_2, \dots, a_n\} \end{array}$$

② Equivalences: A has inverse $\Leftrightarrow$ $\text{rank}(A) = n$ for $A \in \mathbb{C}^{n\times n}$ $\Leftrightarrow$ $\text{range}\,A = \mathbb{C}^n$ $\Leftrightarrow$ $\text{null}\,A = \{0\}$ $\Leftrightarrow$ 0 is not an eigenvalue
$\Leftrightarrow$ 0 is not a singular value
$\Leftrightarrow$ $\det A \neq 0$

③ $(AB)^{-1} = B^{-1}A^{-1}$ $(AB)^* = B^*A^*$

④ Mult. by an orthogonal matrix preserves geometric structure $(Qx)^*(Qy) = x^*$;

⑤ Matrix Norms:

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}| = \text{max abs col sum}$$

$$\|A\|_\infty = \max_{1\le j\le m} \sum_{i=1}^{n} |a_{ij}| = \text{max abs row sum}$$

$$\|A\|_2 = \sqrt{\lambda_{max}(A^*A)} = \sigma_{max}(A)$$

$$\|A\|_F = \left(\sum_{j=1}^{n} \|a_j\|_2^2\right)^{1/2} = \sqrt{tr(A^*A)} = \sqrt{tr(AA^*)}$$; note: $\|AB\|_F^2 \le \|A\|_F^2 \|B\|_F^2$

Theorem: if Q is unitary, $\|QA\|_2 = \|A\|_2$ and $\|QA\|_F = \|A\|_F$

⑥ Absolute Condition Number:

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

$$= \sup_{\delta x} \frac{\|\delta f\|}{\|\delta x\|} \; ; \quad \hat{\kappa} = \|J(x)\|$$ ← norm of Jacobian

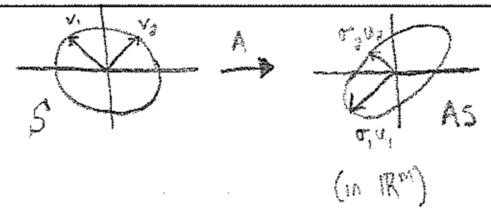
• Relative Condition Number:

$$\kappa = \frac{\|J(x)\|}{\|f(x)\| / \|x\|}$$

• Condition Number of a Matrix: $\kappa(A) = \|A\|\,\|A^{-1}\|$; in 2-Norm: $\kappa(A) = \dfrac{\sigma_1}{\sigma_m}$
(as well as system of eqs.)

Matrix Decompositions

① SVD:



- $\{u_1, \ldots, u_n\}$ are left singular vectors of A in direction of principle axis of AS
- $\{v_1, \ldots, v_n\} \in S$ are right sing. vectors that are preimages of principal semiaxes of AS
- $Av_j = \sigma_j u_j$; $1 \le j \le n$ for $A \in \mathbb{C}^{m \times n}$ w/ $m \ge n$ (A has full rank n)
- $A = \hat{U}\hat{\Sigma}V^*$; U, V are unitary matrices ; $\sigma_k \ge 0 \ \forall k$
- Every matrix has a SVD

① rank of A is # of nonzero sing. values → call it r ∴ range $A = \{u_1, \ldots, u_r\}$
null $A = \{v_{r+1}, \ldots, v_n\}$

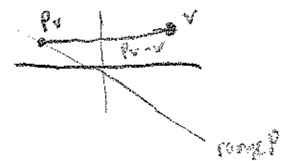
② $\|A\|_2 = \sigma_1$ $\quad \|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_r^2}$

③ Singular values of A are square roots of nonzero eigenvalues of A^*A or AA^*

④ If $A = A^*$, singular values of A |eigenvalues| of A.

⑤ For $A \in \mathbb{C}^{m \times m}$, $|\det A| = \prod_{i=1}^{m} \sigma_i$

② Projectors: $P^2 = P$;



① If $v \in$ range P, $v = Px$ for some x, $Pv = P^2x = Px = v$

$P(Pv - v) = P^2v - Pv = 0$ hence $Pv - v \in$ null P

② Complementary Projectors: If P is a projector, $I - P$ is a comp. proj. to P

If $Pv = 0$, then $(I-P)v = v$ hence $(I-P) \in$ null P

hence $\boxed{\begin{array}{l} \text{range}(I-P) = \text{null}\,P \\ \text{null}(I-P) = \text{range}\,P \end{array}}$; null $(I-P) \cap$ null $P = \{\vec{0}\}$
range $P \cap$ null $P = \{\vec{0}\}$

③ Projector break $\mathbb{C}^m$ into 2 spaces

④ Orthogonal Projectors: iff $P = P^*$ projects onto subspace S_1 along a space S_2

$P = \hat{Q}\hat{Q}^*$ ← orthonormal colums

⑤ Rank 1 Ortho Projs: $P_q = qq^*$ complements are rank $m-1$ orthog. proj. $P_{\perp q} = I - qq^*$

③ QR Factorization: $A = \hat{Q}\hat{R}$

- Assume $A \in \mathbb{C}^{m \times n}$ ($m \ge n$) has full rank n $\qquad \hat{Q}$ has $m \times n$ orthonormal colums
- we need $\langle q_1, \ldots, q_j \rangle = \langle a_1, a_2, \ldots, a_j \rangle$ $j = 1, \ldots, n$ $\qquad \hat{R}$ is $n \times n$ upper Δ
- $a_1 = r_{11} q_1$
  $a_2 = r_{12} q_1 + r_{22} q_2$ $\quad \cdots \quad$ $a_n = r_{1n} q_1 + r_{2n} q_2 + \cdots + r_{nn} q_n$

① Every $A \in \mathbb{C}^{m \times n}$ ($m \ge n$) has a QR factorization (unique reduced $\hat{Q}\hat{R}$ w/ $r_{jj} > 0$)

② Solution of $Ax = b$: suppose $A \in \mathbb{C}^{m \times m}$, $QRx = b \rightarrow Rx = Q^*b$

- Gaussian elimination requires less steps

③ Use Gram-Schmidt to find QR's

↳ classical unstable

④ Gram-Schmidt: $q_1 = \frac{a_1}{r_{11}}$

$q_2 = \frac{a_2 - r_{12} q_1}{r_{22}}$ where $r_{ij} = q_i^* a_j$

$q_n = \frac{a_n - \sum_{i=1}^{n-1} r_{in} q_i}{r_{nn}}$ r_{jj} is taken as $|r_{jj}|$

* unstable b/c of rounding errors *

① Modified Gram-Schmidt: use proj. matrices; $q_n = \frac{P_n a_n}{\|P_n a_n\|}$ where $P_j = I - \hat{Q}_{j-1}\hat{Q}_{j-1}^*$ $\hat{Q}_{j-1} = [q_1 | q_2 | \cdots | q_{j-1}]$

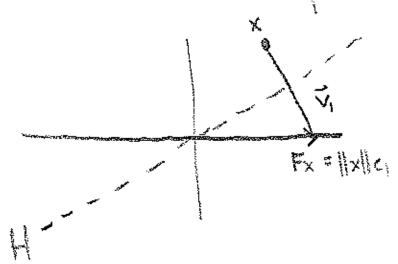
has rank $m-(j-1)$ that projects $\mathbb{C}^m$ onto $\langle q_1, q_2, \ldots, q_{j-1} \rangle^{\perp}$

② Operation Count: $2mn^2$ flops to compute QR factorization

⑤ Householder Triangulation: more stable than Gram-Schmidt : $\underbrace{Q_n Q_{n-1} \cdots Q_2 Q_1}_{Q^*} A = R$ mult. successive unitary matrices

- Gram-Schmidt: Triangular orthogonalization
- Householder: Orthogonal Triangularization
- Operation Count: $2mn^2 - \frac{2}{3}n^3$ flops

Idea: $\underset{A}{\begin{bmatrix} x&x&x\\ x&x&x\\ x&x&x\\ x&x&x \end{bmatrix}} \xrightarrow{Q_1} \underset{Q_1A}{\begin{bmatrix} x&x&x\\ 0&x&x\\ 0&x&x\\ 0&x&x \end{bmatrix}} \xrightarrow{Q_2} \underset{Q_2A}{\begin{bmatrix} x&x&x\\ 0&x&x\\ 0&0&x\\ 0&0&x \end{bmatrix}} \rightarrow \cdots$



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

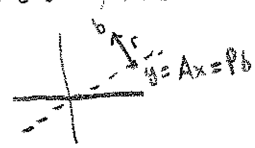
Full rank unitary

$v = \text{sign}(x_1)\|x\| e_1 + x$

where $Q_k = \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix}$ where I is $(k-1)\times(k-1)$ identity

F is $(m-k+1)\times(m-k+1)$ unitary matrix

⑥ Least-Squares Problems: $Ax \approx b$; $A \in \mathbb{R}^{m\times n}$ $(m>n)$ minimize residual; $r = b - Ax \in \mathbb{C}^m$, $x \in \mathbb{C}^n$



① Theorem: r minimizes $\|r\|_2 = \|b - Ax\|_2$ iff $r \perp \text{range}(A)$, $A^* r = 0$

② If A has full rank, then solution is: $x = (A^*A)^{-1}A^*b$ where $A^+ = (A^*A)^{-1}A^*$ is pseudo inverse

Ways to Solve Least Squares:

① Normal Equations:
① Form $A^*A = A^*b$
② Compute Cholesky Factors $A^*A = R^*R$
③ Solve $R^*w = A^*b$ for w then $Rx = w$ for x

Operation Count: $mn^2 + \frac{1}{3}n^3$ - fastest but not always the most stable

② QR Factorization: ① Compute $A = \hat{Q}\hat{R}$
② Compute $\hat{Q}^*b$ and solve $\hat{R}x = \hat{Q}^*b$ for x

$(A^+ = \hat{R}^{-1}\hat{Q}^*)$

Operation count: $2mn^2 - \frac{2}{3}n^3$ flops (using householder)

③ SVD: ① Compute reduced SVD $A = \hat{U}\hat{\Sigma}V^*$
② Compute $\hat{U}^*b$
③ Solve diagonal system $\hat{\Sigma}w = \hat{U}^*b$ for w set $x = Vw$

$A^+ = (V\Sigma^{-1}U^*)$

Operation Count: $2mn^2 + 11n^3$

- use if A is close to rank deficient

Eigenvalue Problems

① Eigenvalue Decompositions: $A = X \Lambda X^{-1}$ ~ change of bases to eigenbases

$$P_A(z) = \det|zI - A|$$

① $A \sim B$ (similar) if $\exists$ a transformation such that $B = X^{-1}AX$

∘ have same eigenvalues, algebraic and geometric multiplicity

Proof: $P_{X^{-1}AX} = P_B = |\lambda I - X^{-1}AX| = |X^{-1}(\lambda I - A)X|$

$= |X^{-1}|\,|\lambda I - A|\,|X|$

$= |\lambda I - A| = P_A$

② Nondefective iff $A = X\Lambda X^{-1}$

③ A Hermitian matrix is unitary diagonalizable (and e-vals are real): $A = Q\Lambda Q^*$; $Q^*Q = QQ^* = I$

∘ Note: Normal matrices $(AA^* = A^*A)$ and are all unitary diagonalizable

② Schur Decompositions: $A = QTQ^*$; Q is unitary T is upper Δ

∘ every square matrix A has a Schur factorization ; $A \sim T$ (eigenvalues of A are on diagonal of T)

∘ if A is normal and we use Schur → it will be same thing as $Q\underset{\uparrow\text{ only diagonal}}{\Lambda}Q^*$

∘ Gershgorin Thm: Every eigenvalue of A lies in at least one of the m circular disks in the complex plane with centers a_{ii} and radii $\sum_{j \neq i} |a_{ij}|$. Moreover if n of these disks form a connected domain that is disjoint from the other $m-n$ disks, then there are n e-values of A within this domain

∘ Eigenvalue Decomps: To compute Schur: Multiply sequence of elementary unitary similarity transforms

$$Q_j^* Q_{j-1}^* \cdots Q_2^* Q_1^* A Q_1 Q_2 Q_3 \cdots Q_j \quad \text{as } j \to \infty \text{ converges to } T$$

Split into 2 Phases:

Ⓘ Direct Method applied to produce an upper Hessenberg matrix H (zeros below 1st sub diagonal)

Ⓘⅈ Generate a sequence of Hessenberg matrices that converge to Δ-form

$$\underset{A \neq A^*}{\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix}} \xrightarrow[\checkmark O(m^3)]{\text{Phase 1}} \underset{H}{\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ & x & x & x \\ & & x & x \end{bmatrix}} \xrightarrow[\sim O(m^3)]{\text{Phase 2}} \underset{T}{\begin{bmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x \end{bmatrix}}$$

(never converges ... usually $O(m)$ to get good approx w/ $O(m^2)$ each step!)

③ Reduction to Hessenberg Form:

$$\underset{A}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix}} \xrightarrow{Q_1^*} \underset{Q_1^*A}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \end{bmatrix}} \xrightarrow{Q_1} \underset{Q_1^*AQ}{\begin{bmatrix} x & & & & \\ x & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix}} \cdots \underset{Q_{m-2}^* Q_{m-1}^* \cdots Q_1^* A Q_1 Q_2 \cdots Q_{m-2}}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & & x & x & x \\ & & & x & x \end{bmatrix}} \quad \text{for } A \in \mathbb{C}^{m\times m}$$

$$\underbrace{Q_{m-2}^* Q_{m-1}^* \cdots Q_1^*}_{Q^*} A \underbrace{Q_1 Q_2 \cdots Q_{m-2}}_{Q}$$

Each Q_k leaves the K^{th} row unchanged

Work for Hessenberg Reduction: $\frac{10}{3}m^3$

For a Hermitian matrix: Hessenberg form is tridiagonal reduction $\sim \frac{4}{3}m^3$

∘ Hessenberg Reduction of $A \in \mathbb{C}^{m\times m}$, $A = QHQ^*$ is backward stable

④ Raylegh Quotient, Inverse Iteration: Assume Real Symmetric Matrices

a) Rayleigh Quotient: $r(x) = \frac{x^T A x}{x^T x}$ $\nabla r(x) = \frac{2}{x^T x}\left(Ax - r(x)x\right)$ → if x is an eigenvector, then $\nabla r = 0$ w/ $r(x)$ = eigenvalue

b) Power iteration: only converges to eigenvector corresponding to the largest e-value

$$v^0 = a_1 q_1 + \dots + a_m q_m$$

$$v^k = c_k A^k v^0 = c_k\left(\lambda_1^k a_1 q_1 + \dots + \lambda_m^k a_m q_m\right) = c_k\left(a_1 q_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^k q_2 + \dots + a_m\left(\frac{\lambda_m}{\lambda_1}\right)^k q_m\right)$$

only linear convergence

c) Inverse Power Iteration:
- For any $\mu \in \mathbb{R}$ the eigenvectors of $(A-\mu I)^{-1}$ are same as A; eigenvalues of $(\lambda_j - \mu)^{-1}$ where $\{\lambda_j\}$ are e-vals of A
- say λ_J is e-value closest to μ → apply power iteration to $(A-\mu I)^{-1}$ to find q_J

* if eigenvalues are known → can use INVERSE ITERATION to find them *

d) Rayleigh Quotient Iteration:

"Use Rayleigh quotient and inverse iteration simultaneously"

$$\text{approx } \lambda_J \xrightarrow{\text{1 step of inverse}} \text{approx } q_J \qquad \text{approx } q_J \xrightarrow{\text{Rayleigh Quotient}} \text{approx } \lambda_J$$

- each iteration triples # of digits of accuracy!

⑤ QR Algorithm w/o Shifts: A is real and symmetric, converges cubically

- will find all eigenvectors if $|\lambda_1| > |\lambda_2| \dots > |\lambda_n| \geq |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \dots \geq |\lambda_m|$
- will find first n eigenvectors of distinct eigenvalues

$$A^k = \underline{Q}^{(k)}\underline{R}^{(k)} \quad \text{now} \quad A^{(k)} = \left(\underline{Q}^{(k)}\right)^T A\, \underline{Q}^{(k)}$$

↓ diagonal elements are Rayleigh Quotients

$A^0 = A$
$A^{(k-1)} = Q^{(k)}R^{(k)}$
$A^{(k)} = R^{(k)}Q^{(k)}$
$\underline{Q}^{(k)} = Q_1^{(1)}Q_2^{(2)} \dots Q^{(k)}$
$\underline{R}^{(k)} = R^{(k)}R^{(k-1)} \dots R^{(1)}$

Theorem: For pure QR applied to real, symmetric matrix, w/ e-values $|\lambda_1| > |\lambda_2| > \dots > |\lambda_m|$ w/ corresponding eigenvector matrix Q has all nonsingular leading principal submatrices. As $k \to \infty$, $A^{(k)}$ converges linearly w/ constant $\max_j \frac{|\lambda_{j+1}|}{|\lambda_j|}$ to diagonal$(\lambda_1, \dots \lambda_m)$ and $\underline{Q}^{(k)}$ converges at the same rate to Q

⑥ QR Algorithm w/ Shifts: uses Rayleigh Quotients to estimate e-values

$$A^{(k-1)} - \mu^{(k)} I = Q^{(k)}R^{(k)}$$

$$\mu^{(k)} = \frac{(q_m^{(k)})^T A\, q_m^{(k)}}{(q_m^{k})^T q_m^{(k)}} = (q_m^{(k)})^T A\, q_m^{(k)} \qquad A^{(k)} = R^{(k)}Q^{(k)} + \mu^{(k)} I$$

$$A^{(k)} = (Q^{(k)})^T A^{(k-1)} Q^{(k)}$$

$$(A - \mu^{(k)} I)(A - \mu^{(k-1)} I)\cdots(A - \mu^{(1)} I) = \underline{Q}^{(k)}\underline{R}^{(k)}$$

* Rayleigh Quotient Shift * $A^{(k)} = (\underline{Q}^{(k)})^T A\, \underline{Q}^{k}$

$\sim \frac{2}{3} m^3$ flops

- backward stable

⑦ Stability of Algorithms: ① Householder Triangularization is backward stable $\forall$ matrices A.

- Solving $Ax = QRx = b$, $Ry = Q^*b$
  $x = R^{-1}y$ (solve $Rx = Q^*b$ by back sub) } all backward stable

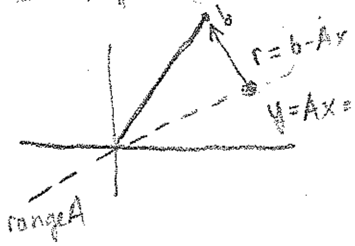
- Backwards Stable: $(A+\Delta A)\tilde{x} = b$ $\quad \frac{\|\Delta A\|}{\|A\|} = O(\epsilon_{mach})$
- Work for back substitution: m^2 for $R \in \mathbb{R}^{m\times m}$

⑧ Conditioning of Least-Squares: Given $A \in \mathbb{C}^{m\times n}$, w/ full rank, $m \geq n$, $b \in \mathbb{C}^m$ find $x \in \mathbb{C}^n$ such that $\|b - Ax\|$ is minimized

Solution: $x = A^+ b = (A^*A)^{-1}A^* b$

$y = P_b = (AA^+)b$

↓ projector onto range A



Theorem: Let $b \in \mathbb{C}^m$, $A \in \mathbb{C}^{m\times n}$ (full rank, $m \geq n$) for least squares, has following 2-norm relative condition numbers describing sensitivities of y and x to perturbations of b and A

	y	x
b	$\frac{1}{\cos\theta}$	$\frac{K(A)}{\eta\cos\theta}$
A	$\frac{K(A)}{\cos\theta}$	$K(A) + \frac{K(A)^2 \tan\theta}{\eta}$

$\eta = \frac{\|A\|\|x\|}{\|y\|}$

note: $K_{b\to y} = \frac{1}{\cos\theta}$ = condition of y w/ respect to perturbations in b

⑨ Stability of Least Squares: Theorems:

① Full-rank least □'s w/ Householder Δ is backward stable (whether $\hat{Q}^*b$ is computed explicitly by forming $\hat{Q}$ first or implicitly)

② Gram-Schmidt for solution of least □'s is backward stable iff $\hat{Q}^*b$ is formed implicitly

③ Solution for least-□'s w/ Normal Equations is unstable in general → can be stable w/ certain restrictions
BUT IS FASTEST

④ Solution of least □'s w/ SVD is backward stable

↳ is stable for Rank-Deficient Least Squares as well!

⑩ Gaussian Elimination – LU Factorization: $A = LU$ ← upper Δ, ↑ lower triangular (1's on diag)

- Work for GE w/ LU: $\frac{2}{m}m^3$
- w/o pivoting: not backward stable
- partial pivoting: $O(m^2)$ extra operations

$Ax = b$ } $\frac{2}{m}m^3$
$LUx = b$
$Ly = b$ } m^2
$Ux = b$ } m^2

∴ approx. half work required by QR

looks like: $\underbrace{L_{m-1}P_{m-1}L_{m-2}P_{m-2}\cdots L_1 P_1}_{L^{-1}} A = U$

BUT: $\underset{\tilde{A}}{(PA)} = LU \longrightarrow \tilde{A} = LU$ ← do Gaussian Elim w/o pivoting

- Gaussian Elimination w/ partial pivoting is backward stable $\quad \frac{\|\delta A\|}{\|A\|} = O(\rho\,\epsilon_{mach})$

growth factor: $\rho = \frac{\max_{ij}|u_{ij}|}{\max_{ij}|a_{ij}|}$

(11) Cholesky Factorization: Hermitian Positive Definite matrices can be decomposed into Δ factors 2x as quickly as other matrices

If A is Hermitian: $A = A^*$, $\lambda \in \mathbb{R}^+$, $v_1 \perp v_2$

Idea: Perform Gaussian elimination:

$$A = \begin{bmatrix} 1 & w^* \\ w & K \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ w & I \end{bmatrix}\begin{bmatrix} 1 & w^* \\ 0 & K - ww^* \end{bmatrix}$$

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

$$\vdots$$

$$= \underbrace{R_1^* R_2^* R_3^* \cdots R_m^*}_{R^*} \underbrace{R_m R_{m-1} \cdots R_1}_{R}$$

$$= R^* R \;;\; R \text{ is upper triangular w/ } r_{ii} > 0$$

- Cholesky Factors are unique.
- Requires $\frac{1}{3}m^3$ operations for
- Always backward stable
- If A is hermitian positive definite, solve $Ax = b$
  - (0) $R^* R x = b$ $\quad \frac{1}{3}m^3$
  - (1) $R^* y = b$ $\quad m^2$
  - (3) $Rx = y$

(12) Review of $Ax = b$: (1) $A \in \mathbb{C}^{m \times n}$ $(m > n)$: use
- QR w/ Householder $\sim \frac{4}{3}m^3 - \ldots$
- SVD if A is rank deficient or close $\sim \frac{4}{3}m^3 + \ldots$
- Normal if all you care about is speed $\sim mn^2 + \frac{1}{3}n^3$ (Form matrix A^*A) then find Cholesky

(2) $A \in \mathbb{C}^{m \times m}$: use LU Factorization in general: $\frac{2}{3}m^3$ flops

(3) $A \in \mathbb{C}^{m \times m}$ and is Hermitian P.D: use Cholesky

(13) Conjugate Gradient: applied to real, positive definite matrices $x^T A x > 0$ ← all positive eigenvalues

- $K_n = \langle b, Ab, \ldots, A^{n-1}b \rangle$, $Ax = b$ exact solution $x_* = A^{-1}$
- minimize A-Norm, $\|x\|_A = \sqrt{x^T A x}$, $\|e_n\|_A = \|x_* - x_n\|_A$
- optimizes $x_n = x_{n-1} + \alpha_n p_{n-1} \rightarrow \phi(x) = \frac{1}{2}x^T A x - x^T b$
- If A has n distinct e-vectors, CG iteration converges at most n steps

$$\frac{\|e_n\|_A}{\|e_0\|_A} \leq 2\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^n \;;\; \kappa = \text{relative condition number}$$