Lecture 19: Numerical Linear Algebra #3
Direct methods for solving Linear Least Squares problems

Outline $A\underline{x}=\underline{b}$

1) Quick Review: Linear Least Squares Problems and $A^TAx=A^Tb$
2) Issues with solving the Normal Equations
3) Orthogonalization, the QR Factorization and Least Squares
   A) Introduction
   B) QR by modified Gram-Schmidt A→Q
   C) QR by Householder Transformation A→R (similarity to LU factorization)

Review: Linear Least-squares problems and the Normal Equations.

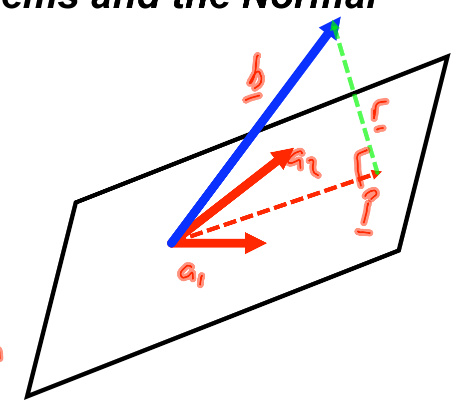
$A\underline{x}=\underline{b}$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \end{array} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \\ \underline{b} \\ \\ \end{bmatrix}$$

full column rank

$\mathrm{rank}(A)=n$

$b \in \mathbb{R}^m$



$A\hat{x}=\underline{p}$ $\quad \underline{r}=\underline{b}-\underline{p}$

$A^TA\hat{x}=A^T\underline{b}$

Normal Equations and Condition number squaring

The SVD and the cond(A)$_2$

$$A = {}_{m}\overset{n}{\begin{bmatrix} & \\ & \\ & \end{bmatrix}} \qquad \text{cond}_2(A) = \frac{\sigma_1}{\sigma_n} \qquad \text{cond}(A^TA) = \frac{\sigma_1^2}{\sigma_n^2} = \text{cond}^2(A)$$

Normal Equations and Condition number squaring

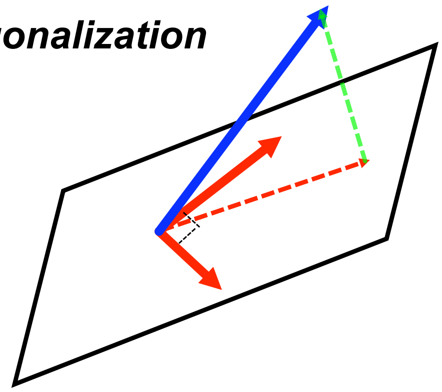
Example: A= [1 a 0; 1 0 a]', b=[1 1 1]

$$A = \begin{bmatrix} 1 & 1 \\ a & 0 \\ 0 & a \end{bmatrix}$$ ← problematic when $a \sim \varepsilon_{mach}$

$$A^TA = \begin{bmatrix} 1 & a & 0 \\ 1 & 0 & a \end{bmatrix} \begin{bmatrix} 1 & 1 \\ a & 0 \\ 0 & a \end{bmatrix} = \begin{bmatrix} 1+a^2 & 1 \\ 1 & 1+a^2 \end{bmatrix}$$ ← problem $a^2 \sim \varepsilon_{mach} \sim 10^{-16}$

$a > \sqrt{\varepsilon_{mach}} \sim 10^{-8}$

Solution of Least Squares by Orthogonalization (QR factorization)



$\underline{A}\,\underline{x} = \underline{b}$

$$A = Q = \begin{array}{c c} & \\ m & \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} \end{array}$$

$\underline{q}_i^T \underline{q}_j = \delta_{ij}$ $\qquad q_i \in \mathbb{R}^m \quad n<m$

$$Q^T Q = \begin{bmatrix} - q_1^T - \\ - q_2^T - \\ - \end{bmatrix} \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_n \\ | & | & | \end{bmatrix} = I$$

Normal Eqs

$A^T A \hat{\underline{x}} = A^T \underline{b}$

$\hat{x} = Q^{-1} \underline{b}$

Solution of Least Squares by Orthogonalization (QR factorization)

Theorem: Every mxn full-column rank Matrix A can be factored as A=QR where Q is mxn and contains an orthonormal basis for C(A) and R is nxn upper-triangular and invertible. (Proof by construction/Gram-Schmidt).

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix} \end{array} = \begin{array}{cc} & n \\ m & \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \\ & Q \end{array} \begin{bmatrix} r_{11} & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & & \cdot \end{bmatrix}$$

diag $\neq 0$

$R \in n \times n$

$C(A) = C(Q)$

Q contains an orthon. basis for C(A)

$\underline{a}_1 = r_{11} \underline{q}_1$

$\underline{a}_2 = r_{21} \underline{q}_1 + r_{22} \underline{q}_2$

Solution of Least Squares by Orthogonalization (QR factorization)

Theorem: if A=QR, the solution of the normal Equations $A^T A\hat{\underline{x}}=A^T\underline{b}$, is the solution of $R\hat{\underline{x}}=Q^T\underline{b}$

$$A^T A\hat{\underline{x}} = A^T\underline{b} \qquad A = QR$$

$$A^T A = R^T Q^T Q R = R^T R$$
$$A^T b = R^T Q^T \underline{b}$$

$$R^T R\hat{x} = R^T Q^T \underline{b}$$
$$R^T(R\hat{x} - Q^T\underline{b}) = \underline{0}$$

if R^T is invertible

$$R\hat{x} = Q^T\underline{b}$$

solve by back subst

QR Factorization Algorithms: #1 Modified Gram-Schmidt Orthogonalization

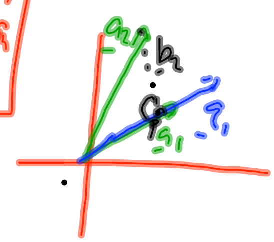
Idea: Take A → Q by repeated projections, get $R=Q^T A$

$$A = QR$$
$$R = Q^T A$$

$$\underline{a}_1, \underline{a}_2, \cdots \underline{a}_n \in \mathbb{R}^m$$

$$\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}$$

$$\underline{q}_1, \underline{q}_2 \cdots q_n \in \mathbb{R}^m$$



$$q_1 = \frac{a_1}{\|a_1\|}$$

$$b_2 = a_2 - \underline{q}_1(\underline{q}_1^T a_2) = (I - q_1 q_1^T)a_2$$

$$(\underline{q}_1 q_1^T)\, a_2$$

$$\underline{q}_2 = \frac{b_2}{\|b_2\|}$$

QR Factorization Algorithms:
#1 Modified Gram-Schmidt Orthogonalization

Example: A= [1 2 2 ; 1 1 0]

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix} \qquad \underline{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \rightarrow \underline{q}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\underline{b}_3 = a_3 - q_1(q_1^T a_3) - q_2(q_2^T a_3)$$

$$q_3 = \frac{b_3}{\|b_3\|}$$

$$\underline{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix} \qquad \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

$$\underline{a}_2 - \underline{q}_1(q_1^T a_2) \rightarrow Q = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \\ 2/3 & -2/3 \end{bmatrix}$$

$$\underline{q}_2 = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

QR Factorization Algorithms:
#1 Modified Gram-Schmidt Orthogonalization

Algorithm:

```
for k=1:n
    r(k,k)=norm(A(:,k))
    A(:,k)=A(:,k)/r(k,k)
    for j=k+1:n
        r(j,k)=A(:,k)'*A(:,j)
        A(:,j)=A(:,j)-r(k,j)*A(:,k)
    END
END
```

$k=1:n$

$$m\begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix}$$

$\underline{q}_1 \; q_2 \; q_3$

$R \begin{bmatrix} & \end{bmatrix}$

$q_k^T a_j$

$a_j = a_j - q_k(q_k^T a_j)$

QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

Idea: Take A → R directly by repeated applications of Q matrices
(similar algorithm to LU factorization)

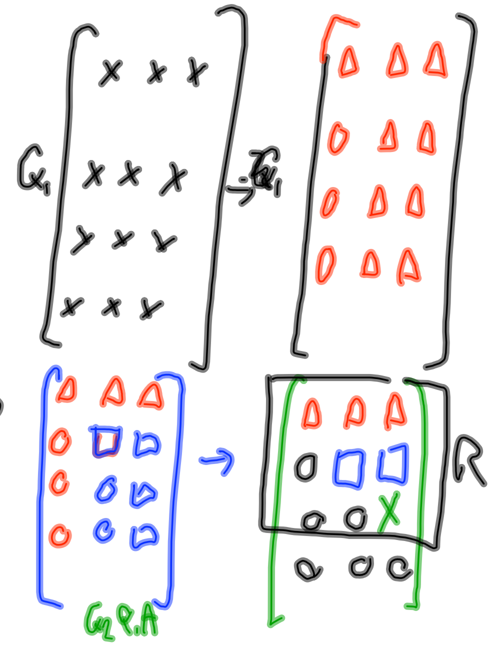
$$E_{n-1}\cdots E_2E_1A = U$$

$$\overbrace{Q_3Q_2Q_1}^{Q^T}A = R$$

$$A = QR$$

$$Q = Q_1^TQ_2^TQ_3^T$$

$$\Rightarrow \quad Rx = Q^Tb$$



QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

The Householder Reflection Matrix $H = I - 2vv^T/v^Tv$ $= I - 2uu^T \quad \underline{u} = \frac{v}{\|v\|}$

Properties:

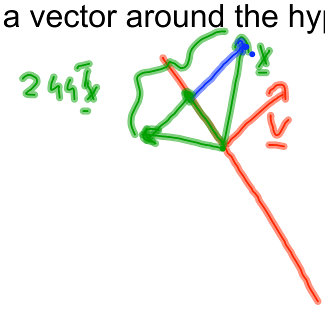
1) Symmetric matrix

$$H^T = H \qquad x' = Q\underline{x} \qquad \|\underline{x}'\| = \|\underline{x}\|$$

2) Q matrix

$$H^TH \overset{?}{=} I = (I-2uu^T)(I-2uu^T) = \underset{=I}{I - 4uu^T + 4u\underbrace{u^Tu}_{}u^T}$$

3) Reflects a vector around the hyperplane orthogonal to $\underline{v}$



$$P = I - uu^T$$

$$y = uu^T\underline{x}$$

QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

Example: Find H to transform a=[1 2 2]' to [c 0 0]'

$$\underline{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \qquad H = I - \frac{2 \underline{v}\underline{v}^T}{\underline{v}^T\underline{v}}$$

$$\|a\| = 3 \qquad H\underline{a} = \begin{bmatrix} \pm\|a\| \\ 0 \\ 0 \end{bmatrix} \qquad a - \frac{2 v v^T a}{v^T v} = \pm\|a\| \underline{e}_1$$

$$\underline{v} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} \qquad c\underline{v} = a - \pm\|a\| \underline{e}_1$$

$$\underline{v} = [\underline{a} + \alpha e_1]$$

$$\alpha = +\text{sign}(a(1))\|\underline{a}\|$$

$$u = \underline{v}/\|\underline{v}\| =$$

$$H = I - \frac{2}{6}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \end{bmatrix}$$

$$H\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{2}{6}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} 6 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix}$$

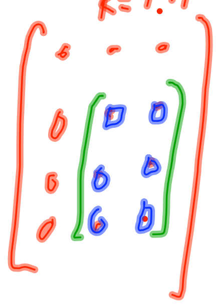
QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

General Idea:

$$Q_1 A$$

$$Q_k = \begin{bmatrix} I_{k-1} & \\ & H_k \end{bmatrix}$$

$k = 1:n$



QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

Example: A= [1 2 2 ; 1 1 0]

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix} \qquad Q_1 = I - \frac{2vv^T}{v^Tv} \qquad \underline{v} = [$$

$$\underline{a}_1 \rightarrow \underline{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \qquad HA$$

$$HA = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix} - \frac{2}{6}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix} - \frac{2}{6}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 6 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \end{bmatrix}$$

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

$$\underline{v}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \text{sign}(0)\,1\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad H_2 = I - \frac{2}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix}$$

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

$$Q_2 Q_1 A = \begin{bmatrix} -3 & -1 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} -3 & -1 \\ 0 & -1 \end{bmatrix}$$

QR Factorization Algorithms:
#2 Orthogonalization by Householder transformation

Algorithm: (quick and dirty)

```
[m,n]=size(A);  % get size of A

for k=1:n   % loop over columns
    i=k:m;   j=k:n; % sub-block indices

    % calculate the reflection vector v
    v=A(i,k); % extract subcolumn k
    v(1)=v(1)+sign(v (1))*norm(v); % modify first component by alpha
    v=v/norm(v); % normalize v

    % Multiply by Qk to transform sublock
    A(i,j) = A(i,j)-2*v*(v'*A(i,j)) ;
end
```

QR Factorization Algorithms: #2 Orthogonalization by Householder transformation

Operation Counts: