Lecture 17. QR decomposition, Orthogonal Matrices and Gram-Schmidt.

For least squares problem $\min_{x \in R^n} \|b - Ax\|^2$

can we use elimination, such as GE to reduce it to tri form, and go from there?

Unfortunately, for a block elimination matrix E

$$\|Er\|^2 \neq \|r\|^2, \quad r = b - Ax, \text{ the residual}$$

Q: What kind of matrices $\underset{\in R^{m\times m}}{Q}$ will statisfy

$$\|Qr\|^2 = \|r\|^2 \quad \forall r \in R^m$$

This requires $r^T Q^T Q r = r^T r \Leftrightarrow r^T \underbrace{(Q^TQ - I)}_{\text{symmetric}} r = 0$

$\forall r \in R^m$, $\Leftrightarrow Q^TQ = I$.

(Th. If $\underset{\in R^{m\times m}}{X} = X^T$, then $r^T X r = 0$ $\forall r \in R^m$ iff. $p^T X q = 0$, $\forall \underset{\in R^m}{p, q}$ iff $X = 0$

Def. $Q \in \mathbb{R}^{m\times m}$ is orthogonal if $Q^TQ = QQ^T = I_m$

Properties of Q: 1) col's of Q independent, and

$$Q = (q_1, \dots, q_m) \quad \begin{cases} q_i^T q_j = 0, & i \neq j \\ q_i^T q_i = 1, & \|q_i\|^2 = 1. \end{cases}$$

The col's of Q form an <u>orthonormal</u> basis of $\mathbb{R}^m$

2) $\forall a \in \mathbb{R}^m$, $\|Qa\| = \|a\|$, orthogonal transforms preserves <u>length</u>.

---

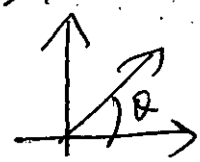
3) $Q^{-1} = Q^T$. $\quad Q = [\overset{n}{Q_1}, Q_2]$

4) orthonormal $\quad Q_1^T Q_1 = I_n$, $n < m$ orthonormal basis.

5) Q_1, Q_2 orth.

Examples. 1) Rotations in $\mathbb{R}^2$ $\quad Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$

$$R(i,j,c,s) \text{ in } \underline{\underline{\mathbb{R}^m}} = \begin{array}{cc} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c \quad 1 \cdots 1 \quad -s & & \\ & & -c \quad \quad s & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow i \\ \leftarrow j \\ \\ \\ \end{array} \end{array}$$



2) Permutation ~~[illegible]~~ $P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, $P^T = P^{-1}$

Permutation + triangular $\Rightarrow$ Identity

Orthogonal + triangular $\Rightarrow$ $\mathrm{diag}(\pm 1, \dots, \pm 1)$.

also of the form $I + uv^T$.

3) Reflection. $H = I - 2uu^T$ with $\|u\| = 1$.

$= I + \underline{(-2u)}\,\underline{u^T}$

The hyperplane $= \{x : u^T x = 0\}$

with normal vector $u \neq 0$.

H is called a Householder matrix.

$H = H^T$, $H^2 = I$, (geometric meaning).

H can be used to do block eliminations: for $a \neq 0$

$$Ha = \|a\| e_1, \quad e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$(I - 2uu^T)a = \|a\| e_1, \quad (2u^T a)u = a - \|a\| e_1$$

If $a \neq \|a\| e_1$, choose $u = \dfrac{a - \|a\| e_1}{\|a - \|a\| e_1\|}$.

$$Ha = \begin{pmatrix} \|a\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(if $a = \|a\| e_1$, also choose $H = I$).

Geometric Significance of Householder Matrices

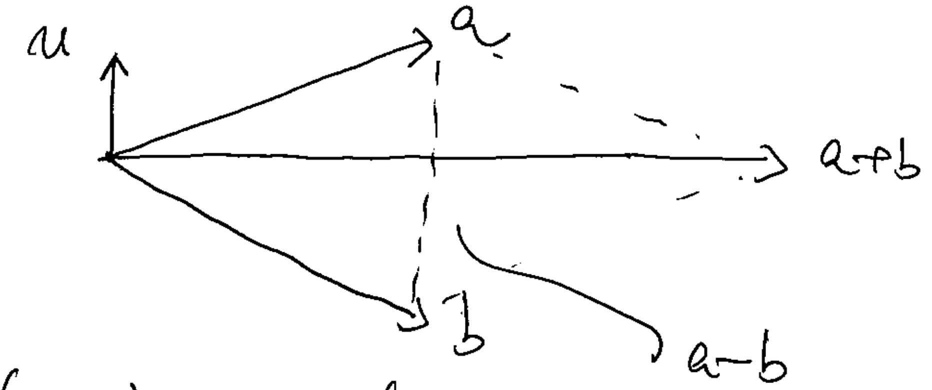
Consider $b = Ha$, where $H = I - 2uu^T$,

and $\|u\| = 1$.

1. $\|b\| = \|Ha\| = \|a\|$, because H is ortho.

2. $b = Ha = a - 2(u^Ta)u$.

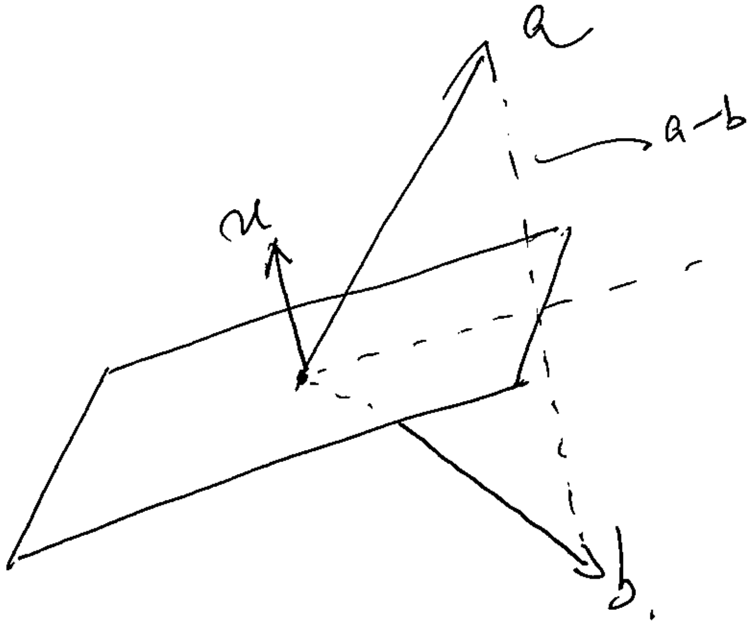
$2(u^Ta)\,u = a - b$. $\Rightarrow$ $u = \pm \dfrac{a-b}{\|a-b\|}$ if $a \neq b$.



Since $(a+b)^T u = \pm (a+b)^T (a-b) / \|a-b\| = 0$.

$a+b$ is in the hyperplane $V = \{ x : x^T u = 0 \}$

V is the bisector of a and b.



The mirror is the hyperplane with normal vector u. The hyperplane is $V=\{x: u^Tx=0\}$

We need to find u, s.t. a and b are mirror images of each other.

Notice that: $(a+b)^T(a-b)=0$ iff $a^Ta=b^Tb$.

i.e. a, b are of the same length.

Now $b = a + b - a = a + \|b-a\|u$, where $u=\frac{b-a}{\|b-a\|}$

$$= a + \|b-a\| u \cdot \frac{u^Ta}{u^Ta} = \left(I + \frac{\|b-a\|}{u^Ta} uu^T\right)a$$

We need to prove $\frac{\|b-a\|}{u^Ta} = -2$. In fact,

$$\frac{\|b-a\|}{u^Ta} = \frac{\|b-a\|^2}{(b-a)^Ta}$$

$$\|b-a\|^2 = b^Tb - 2a^Tb + a^Ta = 2a^Ta - 2a^Tb = 2a^T(a-b)$$

QR decomposition (Reduction via orth. trans to REF).

Th. let $A \in \mathbb{R}^{m\times n}$, $r = \operatorname{rank}(A)$. Then there's orthogonal matrix $Q \in \mathbb{R}^{m\times m}$, s.t.

$$A P = Q \begin{array}{c} \\ r \\ \\ \end{array} \begin{pmatrix} \overset{r}{R_1} & R_2 \\ 0 & 0 \end{pmatrix} \qquad A = \underline{Q}\,\underline{R} \quad R \text{ is REF}$$

where P is a permutation matrix, $R_1 \in \mathbb{R}^{r\times r}$ is upper triangular and nonsingular.

Proof. Use transposition $(1, n_1)$ to ~~ex~~exchange col 1 and the 1st nonzero col of A, if necessary.

$A P(1,n_1) = (a_1, *)$ where $a_1 \neq 0$. and

$P(1,n_1)$ is the transposition matrix.

$a_1 \in R^m$, if $a_1 = \|a_1\| e_1$, ~~just~~ set $H_1 = I_m$

otherwise find Householder matrix H_1 s.t.

$$H_1 a_1 = \|a_1\| e_1 .$$

$$H_1 A p(1, n_1) = \left(\begin{array}{c|c} \|a_1\| & u_1^T \\ \hline 0 & A_1 \end{array}\right),$$ now repeat the same procedure to A_1.

$$\tilde{H}_2 A_1 \tilde{P}(1, n_2) = \left(\begin{array}{c|c} \|a_2\| & * \\ \hline 0 & * \end{array}\right),$$ n_2 is the global index in A.

$$\begin{pmatrix} 1 & \\ & \tilde{H}_2 \end{pmatrix} H_1 A p(1,n_1) \begin{pmatrix} 1 & \\ & \tilde{P}(1,n_2) \end{pmatrix} = \left(\begin{array}{cc|c} \|a_1\| & * & \\ 0 & \|a_2\| & * \\ \hline & 0 & A_2 \end{array}\right)$$

The process will terminate after r steps, since $r(A) = r$.

<u>Remark</u>

1) As in ~~R~~REF, $n_1, n_2, \ldots, n_r$ gives the first r independent col's of A.

2) $Ap = Q_1 (R_1, R_2)$, where $[Q_1, Q_2] = Q$. $Q_1 \in R^{m \times r}$

Span$\{Q_1\} = C(A)$, Q_1 forms an orthonormal basis of $C(A)$, and Q_2 for $N(A^T)$

Now we consider the least squares problem

$$Q^TAP = \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix}.$$

$$\min_{x \in \mathbb{R}^n} \| b - Ax \|^2$$

$$Q^Tb = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad P^Tx = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

We manipulate

$$\| b - Ax \|^2 = \| Q\,(Q^Tb - Q^TAP\,P^Tx) \|^2$$

$$= \left\| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} b_1 - R_1x_1 - R_2x_2 \\ b_2 \end{pmatrix} \right\|^2 = \| b_1 - R_1x_1 - R_2x_2 \|^2 + \|b_2\|$$

$$b_1 - R_1x_1 - R_2x_2 = 0. \Rightarrow x_1 = R_1^{-1}b_1 - R_1^{-1}R_2x_2$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} R_1^{-1}b_1 \\ 0 \end{pmatrix} + \begin{pmatrix} R_1^{-1}R_2 \\ I_{n-r} \end{pmatrix} x_2$$

Solution set:
$$\underbrace{P\begin{pmatrix} R_1^{-1}b_1 \\ 0 \end{pmatrix}}_{\text{special sol.}} + \underbrace{\text{Span}\left\{P\begin{pmatrix} R_1^{-1}R_2 \\ I_{n-r} \end{pmatrix}\right.}_{\text{nullspace}}$$