EE263: Introduction to Linear Dynamical Systems
Review Session 3

Outline

- Orthogonal matrices
- Projection matrices
- QR factorization
- Least squares

Orthogonal matrices

$U \in \mathbf{R}^{n\times n}$ is othogonal $\Longleftrightarrow$

- set of columns is orthonormal
- set of rows is orthonormal

Example: Householder reflections

Householder matrix has form

$$Q = I - 2uu^T$$

where $u^T u = 1$

properties:

- Q is orthogonal
- $Qu = -u$
- $Qv = v$ if $u^T v = 0$

let's verify that Q is orthogonal:

$$\begin{aligned}
Q^TQ &= (I - 2uu^T)^T(I - 2uu^T) \\
&= (I - 2uu^T)(I - 2uu^T) \\
&= I - 2uu^T - 2uu^T + 4uu^Tuu^T \\
&= I - 2uu^T - 2uu^T + 4uu^T \\
&= I
\end{aligned}$$

let's show that $Qu = -u$ and $Qv = v$ when $u^Tv = 0$:

$$\begin{aligned}
Qu &= u - 2uu^Tu = u - 2u = -u \\
Qv &= v - 2uu^Tv = v
\end{aligned}$$

what can we say about a Householder matrix from the fact that

- $Qu = -u$
- $Qv = v$ if $u^T v = 0$?

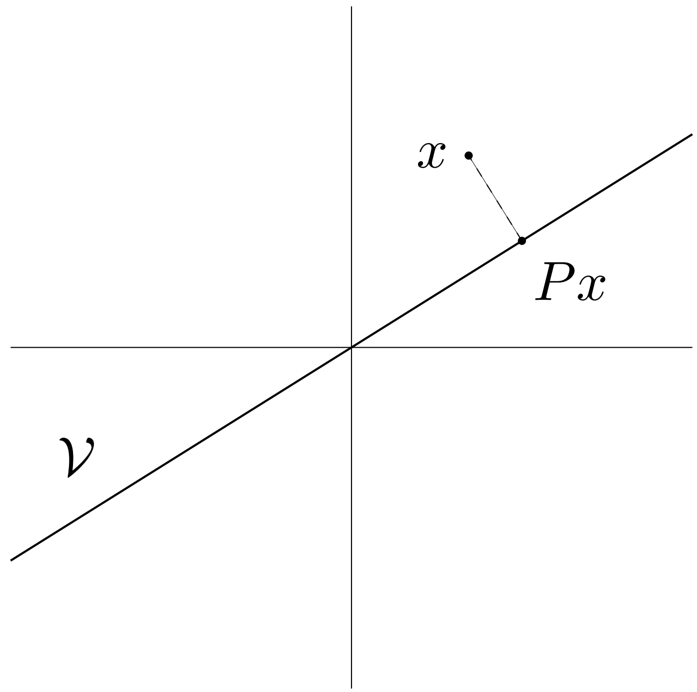
Q represents a reflection through the plane $\{z \mid u^T z = 0\}$ with normal vector u

Projection matrices

a matrix $P \in \mathbf{R}^{n \times n}$ is a *projection matrix* if

- $P = P^T$
- $P^2 = P$

Geometric interpretation



- if we project the projection we obtain the same point
- $P(Px) = Px \to P^2 = P$

you will have to show that if P is a projection matrix so is $I - P$
what kind of projection does $I - P$ represent?

QR factorization

Let $A \in \mathbf{R}^{n \times k}$ with $\mathbf{rank}(A) = r$

- full QR factorization (Lecture 4-20)

$$A = \left[\begin{array}{cc} Q_1 & Q_2 \end{array} \right] \left[\begin{array}{c} R_1 \\ 0 \end{array} \right]$$

where $[Q_1 \; Q_2] \in \mathbf{R}^{n \times n}$ is orthogonal, and $R_1 \in \mathbf{R}^{r \times k}$ is upper staircase

in Matlab

```
>>[Q R] = qr(A);
>>abs(Q*R-A)>1e-10 % check
```

- modified general G-S procedure (Lecture 4-16)

$$A = Q[\tilde{R}\ S]P$$

where $Q^T Q = I_r$, $\tilde{R} \in \mathbf{R}^{r \times r}$ is upper triangular and invertible, and $P \in \mathbf{R}^{k \times k}$ is a permutation matrix

in Matlab

```
>>[Q R E] = qr(A);
>>Q = Q(:,1:r);
>>R_tilde = R(1:r,1:r);
>>S = R(1:r,r+1:end);
>>P = E';
>>abs(Q*[R_tilde S]*P-A)>1e-10 %check
```

'economy' QR factorization in Matlab, *i.e.*, `qr(A,0)`, is **not** equivalent to the above routine

Least-squares

$A \in \mathbf{R}^{m \times n}$ is full rank and skinny

- (typically) no solution of $y = Ax$
- more equations than unknowns

$$\begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \end{bmatrix}$$

$$y \quad = \quad A \quad x$$

What do we do then?

- find the x that's *closest* to satisfying $y = Ax$
- measure of closeness is $\|y - Ax\|$

$$\text{minimize} \quad \|y - Ax\|$$

Often a clever thing to do $\Rightarrow$ very commonly done!

the minimizing x, called the *least-squares (approximate) solution*

$$x_{\rm ls} = A^{\dagger}y = (A^TA)^{-1}A^Ty$$

where $A^{\dagger}$ is a left inverse of A, and is called the pseudo-inverse of A

- Can be found by QR factorization: $A^{\dagger} = R^{-1}Q^T$

in Matlab

```
xls1 = inv(A'*A)*A'*y;
xls2 = A\y;
xls3 = pinv(A)*y;
[Q,R] = qr(A,0);
xls4 = inv(R)*Q'*y;
```

these should all be (very nearly) the same; only numerical roundoff error makes them differ

Uniqueness of left inverse

- is $A^\dagger$ the only left inverse when A is (strictly) skinny and full rank?
- **no!**
- example:
  - consider $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
  - $A^\dagger = [1\ 0]$
  - $B = [1\ 1]$ is another left inverse (since $BA = 1$)

fact: $A^\dagger = (A^T A)^{-1} A^T$ is the *smallest* left inverse of A:

for any B with $BA = I$, we have

$$\|B\|_F = \left(\sum_{i,j} B_{ij}^2 \right)^{1/2} \geq \|A^\dagger\|_F = \left(\sum_{i,j} A_{ij}^{\dagger 2} \right)^{1/2}$$

proof

let $A = QR$ using QR factorization

from $BA = I$, $BQ = R^{-1}$

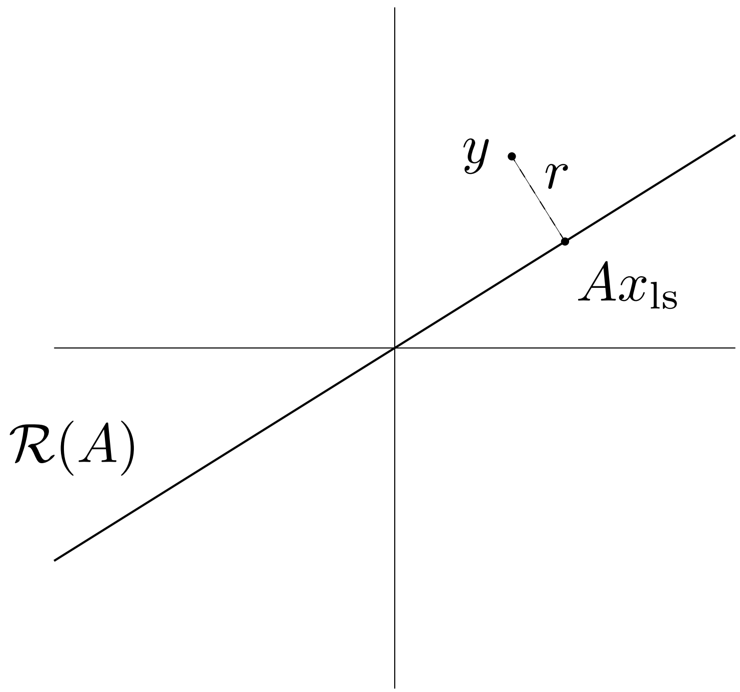
$\|A^\dagger\|_F^2 = \mathbf{Tr}\left(\left((A^T A)^{-1} A^T \right)^T \left((A^T A)^{-1} A^T \right) \right) = \mathbf{Tr}\left((R^T R)^{-1} \right)$

$\|B\|_F^2 \geq \|BQ\|_F^2$

$\|BQ\|_F^2 = \mathbf{Tr}\left((R^T R)^{-1} \right)$

Geometric interpretation

Ax_{ls} is point in $\mathcal{R}(A)$ closest to y (Ax_{ls} is *projection* of y onto $\mathcal{R}(A)$)



Least-squares and projection matrix

from the geometrical intepretation we see

- $Ax_{\rm ls}$ is the projection of y on $\mathcal{R}(A)$
- $Ax_{\rm ls} = A(A^TA)^{-1}A^Ty$
- we expect $A(A^TA)^{-1}A^T$ to be a projection matrix

in fact

$$(A(A^TA)^{-1}A^T)^T = A(A^TA)^{-1}A^T$$
$$A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T$$

Closest point

as we expect we have

$$\|r\|^2 = \|y\|^2 - \|Ax_{\rm ls}\|^2.$$

we first show that $r \perp y_{\rm ls}$
where $y_{\rm ls} = Ax_{\rm ls} = AA^\dagger y = A(A^TA)^{-1}A^Ty$

$$\begin{aligned} {y_{\rm ls}}^T r &= {y_{\rm ls}}^T(y - y_{\rm ls}) = {y_{\rm ls}}^T y \;-\; {y_{\rm ls}}^T y_{\rm ls} \\ &= y^TA(A^TA)^{-1}A^Ty \;-\; y^TA(A^TA)^{-1}A^TA(A^TA)^{-1}A^Ty \\ &= y^TA(A^TA)^{-1}A^Ty \;-\; y^TA(A^TA)^{-1}A^Ty \\ &= 0. \end{aligned}$$

thus, $\|y\|^2 \;=\; \|y_{\rm ls} + r\|^2 = (y_{\rm ls} + r)^T(y_{\rm ls} + r) = \|y_{\rm ls}\|^2 + 2{y_{\rm ls}}^T r + \|r\|^2 = \|y_{\rm ls}\|^2 + \|r\|^2$