Review

Orthogonal Orthonormal

How to convert an arbitrary given set of vectors into an orthogonal / orthonormal set?

$$S = \{u_1, \cdots, u_k\} \xrightarrow[\text{process}]{\text{Gram-Schmidt}} \underset{S' = \{v_1, \cdots, v_k\}}{\text{orthogonal}} \xrightarrow{\text{normalization}} \underset{S'' = \{w_1, \cdots, w_k\}}{\text{orthonormal}}$$

Gram-Schmidt process

$$v_1 = u_1$$

$$v_2 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$\vdots$$

$$v_k = u_k - \frac{u_k \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{u_k \cdot v_{k-1}}{v_{k-1} \cdot v_{k-1}} v_{k-1}$$

normalization $\quad w_1 = \frac{v_1}{\|v_1\|}, \quad w_2 = \frac{v_2}{\|v_2\|}, \quad \cdots, \quad w_k = \frac{v_k}{\|v_k\|}$

$$\mathrm{Span}\{u_1, \cdots, u_k\} = \mathrm{Span}\{v_1, \cdots, v_k\} = \mathrm{Span}\{w_1, \cdots, w_k\}$$

QR-factorization

$m \geq n$

Assume $A \in M_{m\times n}$ with $\mathrm{rank}(A) = n$. Write $A = (a_1 \ a_2 \ \cdots \ a_n)$

$S = \{a_1, \cdots, a_n\} \xrightarrow[\text{process}]{G-S} \{p_1, \cdots, p_n\} \xrightarrow{\text{normalization}} \{q_1, \cdots, q_n\}$ Define $Q = (q_1 \ q_2 \ \cdots \ q_n)$

$\Rightarrow$ a_i is a linear combination of $q_i, q_1, q_2, \cdots, q_{i-1}$ (by G-S process)

$$a_i = (q_1 \ q_2 \ \cdots \ q_n)\begin{pmatrix} r_{1i} \\ \vdots \\ r_{ii} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{with} \quad r_{ii} = \|p_i\| > 0$$

Hence $A = (a_1 \ \cdots \ a_n) = (q_1 \ \cdots \ q_n)\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$

$=: QR$ upper triangular matrix

$\det R = \|p_1\| \cdot \|p_2\| \cdot \cdots \cdot \|p_n\| > 0 \implies R$ is invertible

$Q^T Q = I_n$ (since $\{q_1, q_2, \cdots, q_n\}$ is orthonormal)

Least square approximation

If $Ax = b$ is inconsistent, how to find the "closet" solution?

Find u minimizing $\|Au - b\|$. least square solution

Thm u is a l.s.s. iff u solves for $A^T A x = A^T b$

Proof of existence of solution (in algebra) with rank $A = n$:

Applying the QR-factorization to A,

we find

$$\begin{array}{ccc} A^T A x & = & A^T b \\ \| & & \| \\ R^T Q^T Q R x & & R^T Q^T b \\ \| & & \\ R^T R x & & \end{array} \iff x = R^{-1} Q^T b$$

$\square$

If rank $A \neq n$, singular value decomposition

Finding projection:

① by definition $w_V = \dfrac{w \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \dfrac{w \cdot u_k}{u_k \cdot u_k} u_k$

for any orthogonal basis $\{u_1, \cdots, u_k\}$ of V.

② Suppose $\{u_1, \cdots, u_k\}$ is an orthonormal basis of V,

$Q := (u_1 \cdots u_k)$. then $w_V = Q Q^T w$

③ Suppose $\{u_1, \cdots, u_k\}$ is an arbitrary basis of V.

$A := (u_1 \cdots u_k)$, then $w_V = A (A^T A)^{-1} A^T w$

proof: ① $\iff$ ② $\quad w_V = Q Q^T w = (u_1 \cdots u_k) \begin{pmatrix} u_1^T \\ \vdots \\ u_k^T \end{pmatrix} w = (u_1 \cdots u_k) \begin{pmatrix} u_1 \cdot w \\ \vdots \\ u_k \cdot w \end{pmatrix}$

② $\iff$ ③ since rank $A = k$, $A = QR$ (QR-factorization)

$w_V = A (A^T A)^{-1} A^T w = QR(R^T \underbrace{Q^T Q}_{I} R)^{-1} R^T Q^T w = Q \underbrace{R R^{-1} (R^T)^{-1} R^T}_{I} Q^T w$ $\square$

Geometric explanation of linear systems (2^{nd} Cont.)

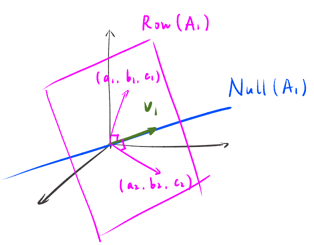
Motivation to
Least square approximation

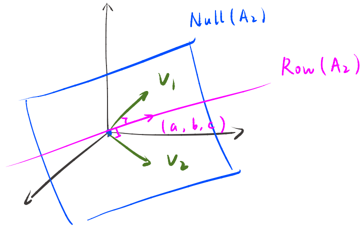
Consider the inhomogeneous linear system $Ax = b$.

In tutorial 4, we see that

if consistent, then $\underbrace{\overbrace{\{x : Ax = b\}}^{\text{affine space}}}_{\substack{\text{general solutions} \\ \text{to } Ax=b}} = \underbrace{u}_{\substack{\text{one solution} \\ \text{to } Ax=b}} + \underbrace{\overbrace{\{y : Ay = 0\}}^{\text{vector space}}}_{\substack{\text{general solutions} \\ \text{to } Ax=0,\ \text{or } Null(A)}}$ translation along u

In Tutorial 7, we give the geometric pictures of $Null(A)$, $Row(A)$:





$$A_1 = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

$$\text{or } = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \tilde{a}_3 & \tilde{b}_3 & \tilde{c}_3 \end{pmatrix}$$

rank $A_1 = 2$

$(\tilde{a}_3\ \tilde{b}_3\ \tilde{c}_3)$ is a linear combination of $\{(a_1\ b_1\ c_1), (a_2\ b_2\ c_2)\}$

$$A_2 = (a\ \ b\ \ c)$$

$$\text{or } = \begin{pmatrix} a & b & c \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{c}_1 \\ \tilde{a}_2 & \tilde{b}_2 & \tilde{c}_2 \end{pmatrix}$$

rank $A_2 = 1$

$(\tilde{a}_1\ \tilde{b}_1\ \tilde{c}_1), (\tilde{a}_2\ \tilde{b}_2\ \tilde{c}_2)$ are some multiple of $(a\ b\ c)$

Aim: characterize $Null(A, b) := \{x : Ax = b\}$

Assuming $Null(A, b) \neq \phi$, namely $Ax = b$ is consistent,

we have an observation:

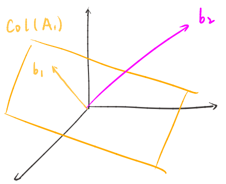
$A = (c_1 \cdots c_n)$

consistent iff $\exists \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is a solution to $Ax = b$.

iff $b = (c_1 \cdots c_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 c_1 + \cdots + x_n c_n \in Col(A)$

$$A_1 = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \tilde{a}_3 & \tilde{b}_3 & \tilde{c}_3 \end{pmatrix}$$

rank $A_1 = 2$



b_1 consistent

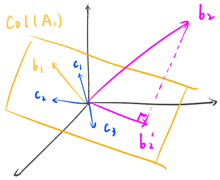
b_2 inconsistent

$\longrightarrow$ solvability : $Col(A) \subseteq \mathbb{R}^m$

solutions : $Row(A)$ & $Null(A) \subseteq \mathbb{R}^n$

Assuming $Null(A, b) = \phi$, namely $Ax = b$ is inconsistent,

we want to find an approximate solution (least square approximation)



Let b_2' be the projection of b_2 onto $col(A)$.

Then $Ax = b_2'$ is consistent.

with solutions $(x_1, \cdots, x_n)^T$ satisfying

$$b_2' = x_1 c_1 + \cdots + x_n c_n$$

least square solution

Remark :

The vector b_2' is unique,

but the l.s.s. might be not unique.

L.S.S is unique iff $\{c_1, \cdots, c_n\}$ is linearly independent

iff $\operatorname{rank} A = n$

1. Apply Gram-Schmidt Process to convert

(a) $$\begin{array}{cccc} u_1 & u_2 & u_3 & u_4 \\ \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, & \begin{pmatrix}1\\-1\\1\\0\end{pmatrix}, & \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}, & \begin{pmatrix}1\\2\\0\\1\end{pmatrix}\end{array}$$ into an orthonormal basis for $\mathbb{R}^4$.

(b) $$\left\{\begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\1\end{pmatrix}\right\}$$ into an orthonormal set. Is the set obtained an orthonormal basis? Why?

1. (a) orthogonal (G-S process)

$$v_1 = u_1$$

$$v_2 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \frac{1}{4}(3, -5, 3, -1)^T$$

$$v_3 = u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 = \frac{2}{11}(7, 3, -4, -6)^T$$

$$v_4 = u_4 - \frac{u_4 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_4 \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{u_4 \cdot v_3}{v_3 \cdot v_3} v_3 = \frac{1}{10}(1, -1, -2, 2)^T$$

orthonormal (normalization)

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{2}(1, 1, 1, 1)^T$$

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{1}{2\sqrt{11}}(3, -5, 3, -1)^T$$

$$w_3 = \frac{v_3}{\|v_3\|} = \frac{1}{\sqrt{110}}(7, 3, -4, -6)^T$$

$$w_4 = \frac{v_4}{\|v_4\|} = \frac{1}{\sqrt{10}}(1, -1, -2, 2)^T$$

(b) similarly compute $v_1 = u_1 \qquad v_2 = \frac{1}{10}(3, 6, -4, -7)^T$

$$v_3 = \frac{2}{11}(4, -3, 2, -2)^T \qquad v_4 = 0$$

$\Longrightarrow$ Not a basis. because these vectors are not linearly independent

(or the projection of v_4 onto $\operatorname{span}\{v_1, v_2, v_3\}$ is zero)

$$u_4 = \frac{u_4 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{u_4 \cdot v_2}{v_2 \cdot v_2} v_2 + \frac{u_4 \cdot v_3}{v_3 \cdot v_3} v_3$$

$$= \underbrace{\cdots}_{\text{numbers}} u_1 + \underline{\cdots}\, u_2 + \underline{\cdots}\, u_3$$

2. Let $\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 6 \\ 3 \\ -1 \\ 1 \end{pmatrix}$.

(a) Is the linear system $\mathbf{Ax} = \mathbf{b}$ inconsistent?

(b) Find a least squares solution to the system. Is the solution unique?

(c) Use your answer in (b), compute the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$. Is the solution unique?

2. (a) Yes. $\operatorname{rref}(A \quad b) = \left(\begin{array}{cccc|c} 1 & & & 1 & \\ & 1 & & 1 & \\ & & 1 & -1 & \\ & & & & 1 \end{array}\right)$

(b) Solve for $(A^T A)x = A^T b \Longrightarrow \begin{cases} x_1 = -6 - s \\ x_2 = -1 - s \\ x_3 = 7 + s \\ x_4 = s \end{cases} \quad s \in \mathbb{R}$

Take $s = 0$ $\begin{pmatrix} -6 \\ -1 \\ 7 \\ 0 \end{pmatrix}$

(c) The projection of b onto $col(A)$ is $Av = (6, 2, 1, 0)^T$.

Yes. The projection is unique. Moreover we can check

$$A \begin{pmatrix} -6-s \\ -1-s \\ 7+s \\ s \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad \forall s \in \mathbb{R}$$

3. **(Application)** A line

$$p(x) = a_1 x + a_0$$

is said to be the *least squares approximating line* for a given a set of data points (x_1, y_1), (x_2, y_2), ..., (x_m, y_m) if the sum

$$S = [y_1 - p(x_1)]^2 + [y_2 - p(x_2)]^2 + \cdots + [y_m - p(x_m)]^2$$

is minimized. Writing

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}, \text{ and } p(\mathbf{x}) = \begin{pmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_m) \end{pmatrix} = \begin{pmatrix} a_1x_1 + a_0 \\ a_1x_2 + a_0 \\ \vdots \\ a_1x_m + a_0 \end{pmatrix}$$

the problem is now rephrased as finding a_0, a_1 such that

$$S = ||\mathbf{y} - p(\mathbf{x})||^2$$

is minimized. Observe that if we let

$$\mathbf{N} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \text{ and } \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix},$$

then $\mathbf{Na} = p(\mathbf{x})$. And so our aim is to find $\mathbf{a}$ that minimizes $||\mathbf{y} - \mathbf{Na}||^2$.

It is known the equation representing the dependency of the resistance of a cylindrically shaped conductor (a wire) at 20^oC is given by

$$R = \rho \frac{L}{A},$$

where R is the resistance measured in Ohms Ω, L is the length of the material in meters m, A is the cross-sectional area of the material in meter squared m^2, and ρ is the resistivity of the material in Ohm meters Ωm. A student wants to measure the resistivity of a certain material. Keeping the cross-sectional area constant at $0.002m^2$, he connected the power sources along the material at varies length and measured the resistance and obtained the following data.

L	0.01	0.012	0.015	0.02
R	2.75×10^{-4}	3.31×10^{-4}	3.92×10^{-4}	4.95×10^{-4}

It is known that the Ohm meter might not be calibrated. Taking that into account, the student wants to find a linear graph $R = \frac{\rho}{0.002}L + R_0$ from the data obtained to compute the resistivity of the material.

(a) Relabeling, we let $R = y$, $\frac{\rho}{0.002} = a_1$ and $R_0 = a_0$. Is it possible to find a graph $y = a_1 x + a_0$ satisfying the points?

(b) Find the least square approximating line for the data points and hence find the resistivity of the material. Would this material make a good wire?

3. (a)

$$A = \begin{pmatrix} 1 & 0.01 \\ 1 & 0.012 \\ 1 & 0.015 \\ 1 & 0.02 \end{pmatrix} \quad b = \begin{pmatrix} 2.75 \times 10^{-4} \\ 3.31 \times 10^{-4} \\ 3.92 \times 10^{-4} \\ 4.95 \times 10^{-4} \end{pmatrix} \quad Ax = b \text{ inconsistent}$$

(b) Solve for $(A^TA)x = A^Tb \Rightarrow x = \begin{pmatrix} 0.0001 \\ 0.0216 \end{pmatrix}$

$\Rightarrow \rho = 0.002 \times 0.0216 = 4.32 \times 10^{-5}\ \Omega m$

4. **(Application)** Suppose the equation governing the relation between data pairs is not known. We may want to then find a polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

of degree n, $n \le m-1$, that best approximates the data pairs (x_1, y_1), (x_2, y_2), ..., (x_m, y_m). A *least square approximating polynomial* of degree n is such that

$$||\mathbf{y} - p(\mathbf{x})||^2$$

is minimized. If we write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}, \mathbf{N} = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix} \text{ and } \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix},$$

then $p(\mathbf{x}) = \mathbf{Na}$, and the task is to find $\mathbf{a}$ such that $||\mathbf{y} - \mathbf{Na}||^2$ is minimized. Observe that $\mathbf{N}$ is a matrix minor of the Vandermonde matrix. If at least $n+1$ of the x-values $x_1, x_2, ..., x_m$ are distinct, the columns of $\mathbf{N}$ are linearly independent, and thus $\mathbf{a}$ is uniquely determined by

$$\mathbf{a} = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{y}.$$

We shall now find a quartic polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

that is a least square approximating polynomial for the following data points

x	4	4.5	5	5.5	6	6.5	7	8	8.5
y	0.8651	0.4828	2.590	-4.389	-7.858	3.103	7.456	0.0965	4.326

Enter the data points.

```
>> x=[4 4.5 5 5.5 6 6.5 7 8 8.5]';
```

```
>> y=[0.8651 0.4828 2.590 -4.389 -7.858 3.103 7.456 0.0965 4.326]';
```

Next, we will generate the 10×10 Vandermonde matrix.

```
>> N=fliplr(vander(x));
```

We only want the matrix minor up to the 4-th power, that is, up to the the 5-th column,

```
>> N=N(:,1:5);
```

Use this to find the least square approximating polynomial of degree 4.

4. a = inv(N' * N) * N' * y

$$\text{ans} = \begin{pmatrix} -204.0716 \\ 169.2099 \\ -49.7013 \\ 6.1528 \\ -0.2720 \end{pmatrix}$$

$y = -0.2720x^4 + 6.1528x^3 - 49.7013x^2 + 169.2099x - 204.0716.$

5. Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

$A = (c_1 \ \ c_2 \ \ c_3)$

(a) Find a QR factorization of $\mathbf{A}$.

(b) Use your answer in (a) to find the least square solution to $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

5. (a) G-S process $\Rightarrow$

$$u_1 = c_1 = (1, 1, 1, 0)^T$$

$$u_2 = c_2 - \frac{c_2 \cdot u_1}{u_1 \cdot u_1} u_1 = (0, 0, 0, 1)^T$$

$$u_3 = c_3 - \frac{c_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{c_3 \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{1}{3}(-1, -1, 2, 0)^T$$

normalization $\Rightarrow$

$$v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{3}}(1, 1, 1, 0)^T$$

$$v_2 = \frac{u_2}{\|u_2\|} = (0, 0, 0, 1)^T$$

$$v_3 = \frac{u_3}{\|u_3\|} = \frac{1}{\sqrt{6}}(-1, -1, 2, 0)^T$$

So $$Q = (v_1 \ \ v_2 \ \ v_3) = \begin{pmatrix} 1/\sqrt{3} & 0 & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 0 & 1 & 0 \end{pmatrix}$$

$$R = Q^T A = \begin{pmatrix} \sqrt{3} & \sqrt{3} & 1/\sqrt{3} \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{6}/3 \end{pmatrix}$$

(b)

$$(A^T A)x = A^T b \iff x = R^{-1} Q^T b = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

$$\begin{array}{ccc} (A^T A)\, x & = & A^T b \\ \| & & \| \\ (R^T Q^T Q R)\, x & & R^T Q^T b \\ \| & & \\ (R^T R)\, x & & \end{array}$$