EE5120 Linear Algebra: Tutorial 4, July-Dec 2018, Dr. Uday Khankhoje, EE IIT Madras

Covers Ch 3.1,3.2,3.3,3.4 of GS

1. Given a matrix P that satisfies $P^2 = P$ and $P^T = P$. Using these facts, prove that Pb is the projection of b onto the column space of P.

   Hint: Use appropriate orthogonality relations of subspaces

> **Solution:** (added by Uday)
> Let b be a point to be projected, and c be the projected point, i.e. $c = Pb$. If P is a projection operator, the difference vector between the point and its projection $(c - b)$ should not have any component along the projection $(Px\,\forall x)$, else the entire operation would not have been a projection; mathematically, this reads as: $(c-b)^T(Px) = 0$. Simplifying, $(Pb - b)^T(Px) = b^TP^TPx - b^TPx$. If we impose the given conditions $P^2 = P$ and $P^T = P$, we get $(c-b)^T(Px) = 0$. QED

2. (a) Consider the system of linear equations $A\mathbf{x} = \mathbf{b}$, where A is a full column rank matrix, $\mathbf{x}$ is an n-length vector and $\mathbf{b}$ is an m-length vector. What is the least squares (LS) solution to the above system? Prove that the error in the estimate is in the left null-space of A.

   (b) Once again consider $A\mathbf{x} = \mathbf{b}$, with $A = \begin{bmatrix} 1 & 3 & -1 & 3 & 2 \\ 2 & -1 & 0 & 1 & 0 \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} -1 & 4 & 1 & 2 & 1 \end{bmatrix}^T$. Can you find a solution to the given system of linear equations using Gauss elimination? Can you find a LS solution? Verify (a) for this example. Further, determine the left null space of A and verify actually whether the LS error lies in the left null space of A or not, by expressing the error in terms of the basis of left null space of A.

> **Solution: Added by Manoj**
>
> (a) LS solution is given by $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$. Error is given by $\mathbf{e} = A\hat{\mathbf{x}} - \mathbf{b}$. Now, we have,
>
> $$A^T\mathbf{e} = A^T\Big(A\hat{\mathbf{x}} - \mathbf{b}\Big) = A^T\Big(A(A^TA)^{-1}A^T\mathbf{b} - \mathbf{b}\Big) = A^TA(A^TA)^{-1}A^T\mathbf{b} - A^T\mathbf{b}$$
> $$= (I)A^T\mathbf{b} - A^T\mathbf{b} = 0.$$
>
> Hence, the LS error lies in the left null space of A.
>
> (b) Echelon form of $[A|\mathbf{b}]$ result in the matrix,
>
> $$\begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
>
> which implies **solution via Gauss elimination cannot be found**. This is because given $\mathbf{b}$ doesn't lie in the column space of A. The LS solution to the given problem is $\hat{x} = [0.8286 \quad -0.9429]^T$ and the error is
>
> $$\begin{bmatrix} -0.0571 & -0.5714 & -1.8286 & -0.4571 & 0.6571 \end{bmatrix}^T.$$

The left null space matrix for A is

$$L_A = \begin{bmatrix} \frac{1}{7} & -\frac{6}{7} & -\frac{2}{7} \\ \frac{2}{7} & -\frac{5}{7} & -\frac{4}{7} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Further, it can be verified that row reduced echelon form of $[L_A \big| \mathbf{e}]$ is,

$$\begin{bmatrix} 1 & 0 & 0 & -1.8286 \\ 0 & 1 & 0 & -0.4571 \\ 0 & 0 & 1 & 0.6571 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It implies the error lies in the column space of L_A, i.e., it can be expressed in terms of basis of left null space of A. This **also** ensures that LS error lies in left null space of A.

3. Consider the QR decomposition of a matrix A as shown below. Matrix on the LHS is A. On the RHS, first matrix is Q and is multiplied with R. Fill in the blanks.

$$\begin{bmatrix} 1 & \text{----} & -1 & \text{----} \\ 1 & \text{----} & 2 & \text{----} \\ -1 & \text{----} & 3 & \text{----} \\ 1 & \text{----} & 1 & \text{----} \end{bmatrix} = \begin{bmatrix} \text{----} & 0 & \text{----} & \frac{6}{\sqrt{72}} \\ \text{----} & \frac{1}{\sqrt{2}} & \text{----} & -\frac{4}{\sqrt{72}} \\ \text{----} & \frac{1}{\sqrt{2}} & \text{----} & \frac{4}{\sqrt{72}} \\ \text{----} & 0 & \text{----} & \frac{2}{\sqrt{72}} \end{bmatrix} \begin{bmatrix} \text{----} & 0 & \text{----} & 0.5 \\ \text{----} & \sqrt{2} & \text{----} & 2.1213 \\ \text{----} & 0 & \text{----} & 0.5 \\ \text{----} & 0 & \text{----} & 1.4815 \end{bmatrix}.$$

Solution: Added by Manoj.

- First column in Q is simply the first column of A normalized to unit norm. Hence, it is given by $\frac{1}{2}[1\ 1\ \ -1\ 1]^T$.
- Given first two columns of Q and the third column of A, 3rd column of Q can be computed from Gram-Schmidt. It is $\frac{1}{6}[-3\ \ -1\ 1\ 5]^T$.
- Recall R is a upper-triangular matrix. So, in the first column of R, first entry is the inner product between 1st columns of A and Q, and remaining entries in first column of R are zeros. Hence, it can be verified that the 1st column of R is $[2\ 0\ 0\ 0]^T$.
- 3rd column of R is computed as follows: first 3 entries are the inner products between first 3 columns of Q with the 3rd column of A. So, 3rd column of R will be $[-0.5, 3.5355, 1.5, 0]^T$.
- Since, entire Q and R matrices are now obtained, on multiplying them, we get the second and fourth columns of A which are $[0, 1, 1, 0]^T$ and $[1.04, 0.968, 2.03, 1.015]^T$ respectively.

4. (a) Prove that the trace of $P = \mathbf{a}\mathbf{a}^T/\mathbf{a}^T\mathbf{a}$—which is the sum of its diagonal entries—- always equal 1.
   (b) Is the projection matrix P invertible? Why or why not?

Solution: (Added by Yaswanth)

(a) Let $\mathbf{a}$ be of size $n \times 1$. Then,

$$\mathbf{a^T a} = \sum_{i=1}^{n} \mathbf{a_i^2}$$

Let A= $\mathbf{aa^T}$. Then, i^{th} diagonal elements of A is A_{ii}= $\mathbf{a_i^2}$

$$\implies \text{Sum of Diagonal elements of } \mathbf{aa^T} = \sum_{i=1}^{n} \mathbf{a_i^2}$$

$$\implies \text{Trace of } \frac{\mathbf{a^T a}}{\mathbf{aa^T}} = 1.$$

(b) Projection operator is projects any vector on to a subspace. So, The dimension of subspace is always less than or equal to the vectorspace, to which it belongs. So, Projection operator is not invertible.
(OR)
We Know, $P^2 = P$.
Let us assume the P is invertible,i.e., $P^{-1}P = I$

$$\implies P^2 = IP = P^{-1}PP = P^{-1}P = I$$

$$\implies P = I$$

So, Identity matrix is the only projection matrix which has inverse.

5. (a) Show that length of $A\mathbf{x}$ equals the length of $A^T\mathbf{x}$ iff $AA^T = A^TA$.
   (b) If Q_1 and Q_2 are orthogonal matrices in $\mathbb{R}^{2\times 2}$, so that $Q^TQ = I$, show that Q_1Q_2 is also orthogonal. If Q_1 is rotation through θ, and Q_2 is rotation through ϕ, what is Q_1Q_2? Can you find the trigonometric identities for $\sin(\theta + \phi)$ and $\cos(\theta + \phi)$ in the matrix multiplication Q_1Q_2?

Solution: (Added by Yaswanth)

(a) Given $||A\mathbf{x}|| = ||A^T\mathbf{x}||$

$$\implies ||A\mathbf{x}||^2 = ||A^T\mathbf{x}||^2$$

$$\implies (A\mathbf{x})^T(A\mathbf{x}) = (A^T\mathbf{x})^T(A^T\mathbf{x})$$

$$\implies \mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^TAA^T\mathbf{x}$$

$$\implies \mathbf{x}^T(A^TA - AA^T)\mathbf{x} = 0$$

This should satisfy for all $\mathbf{x}$,
So, $A^TA = AA^T$. Similarly, if $A^TA = AA^T$, then $||A\mathbf{x}||^2 = \mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^TAA^T\mathbf{x} = ||A^T\mathbf{x}||^2$. Hence, proved.

(b) Given Q_1, Q_2 are orthogonal. So, $Q_1^T Q_1 = I, Q_2^T Q_2 = I$.
Then, $(Q_1Q_2)^T(Q_1Q_2) = Q_2^T(Q_1^TQ_1)Q_2 = Q_2^TQ_2 = I$. So, Q_1Q_2 is also orthogonal matrix.
Q_1, Q_2 are rotations through θ, ϕ respectively. Then,

$$Q_1 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, Q_2 = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

$$\implies Q_1Q_2 = \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix}$$

$$\implies Q_1Q_2 = \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}$$

6. (a) Teacher taught kinematics in the class that, $s = ut + \frac{1}{2}at^2$. She asked all the students to find out what is the earth's acceleration due to gravity by doing a experiment at home. What would you do to get an accurate value.

(b) Given a system with input and output relation: $y = ax + bx^2$. To find the coefficients, an experiment is done and data is shown below. Find the coefficients?

x	20	21	22	23	24	25	26
y	2396	2562	2727	2904	3070	3254	3427

Hint: Use Least square error curve fit.

Solution: (Added by Yaswanth)

(a) There are many ways. One of the ways is, go to terrace of your house, drop stone with initial velocity zero m/s and measure the time taken for reaching the ground. Repeat the experiment multiple times. Now measure the height of your house. Then, use $s = \frac{1}{2}at^2$ and solve for 'a' having least square error.

(b)

$$\begin{bmatrix} 20 & 400 \\ 21 & 441 \\ 22 & 484 \\ 23 & 529 \\ 24 & 576 \\ 25 & 625 \\ 26 & 676 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2396 \\ 2562 \\ 2727 \\ 2904 \\ 3070 \\ 3254 \\ 3427 \end{bmatrix}$$

Above equation is of the form $A\mathbf{z} = \mathbf{g}$. Solving for Least Square error, $A^TA\mathbf{z} = A^T\mathbf{g}$

$$\begin{bmatrix} 3731 & 87101 \\ 87101 & 2047955 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 472640 \\ 11063048 \end{bmatrix}$$

Solving, $a = 79.927, b = 2.003$. Approximately, $a = 80, b = 2$, i.e., $y = 80x + 2x^2$.

7. Let $\mathbb{P}_3(t)$ be the vector space of polynomials at degree at most equal to 3, defined with

inner product $\Big(f(t), g(t)\Big) = \int_{-1}^{1} f(t)g(t)dt$, for any two functions $f(t), g(t) \in \mathbb{P}_3$. Apply the Gram-Schmidt procedure to $\{1, t, t^2, t^3\}$ to find an orthonormal basis $\{f_0, f_1, f_2, f_3\}$.

Solution: (Added by Prajosh)
Consider, $f(t)$ and $g(t)$ are t^r and t^s respectively and take $r + s = n$. Then,

$$
\begin{aligned}
< t^r, t^s > &= \int_{-1}^{1} f(t)g(t)dt = \frac{t^{n+1}}{n+1}\Big|_{-1}^{1} \\
&= \begin{cases} \frac{2}{n+1}, & \text{if } n \text{ is even,} \\ 0, & \text{else.} \end{cases}
\end{aligned}
$$

Now, for the given problem, we have the following:

- Set $\tilde{f}_0 = 1$. Normalizing to unit length we have $f_0 = \frac{1}{\sqrt{2}}$.
- Compute $\tilde{f}_1 = t - \frac{<t, f_0>}{<f_0, f_0>}(f_0) = t - 0 = t$. So, we get f_1, on normalizing $\tilde{f}_1$ to unit length. Hence, $f_1 = \sqrt{\frac{3}{2}}t$.
- Compute,

$$
\begin{aligned}
\tilde{f}_2 &= t^2 - \frac{< t^2, f_0 >}{< f_0, f_0 >}(f_0) - \frac{< t^2, f_1 >}{< f_1, f_1 >}(f_1) \\
&= t^2 - \frac{1}{\sqrt{2}}.\frac{2}{3}.\frac{1}{\sqrt{2}} - \sqrt{\frac{3}{2}}.0.(\sqrt{\frac{3}{2}}t \\
&= t^2 - \frac{1}{3}.
\end{aligned}
$$

  Thus, $f_2 = \sqrt{\frac{45}{8}}\left(t^2 - \frac{1}{3}\right)$.

- Compute,

$$
\begin{aligned}
\tilde{f}_3 &= t^3 - \frac{< t^3, f_0 >}{< f_0, f_0 >}(f_0) - \frac{< t^3, f_1 >}{< f_1, f_1 >}(f_1) - \frac{< t^3, f_2 >}{< f_2, f_2 >}(f_2) \\
&= t^3 - 0 - \sqrt{\frac{3}{2}}.\frac{2}{5}.\sqrt{\frac{3}{2}}t - 0 \\
&= t^3 - \frac{3}{5}t.
\end{aligned}
$$

  Hence, $f_3 = \sqrt{\frac{175}{8}}\left(t^3 - \frac{3}{5}t\right)$.

8. Let $\bar{x} = \frac{1}{n}(x_1 + ... + x_n)$ and $\bar{y} = \frac{1}{n}(y_1 + ... + y_n)$. Show that the least-squares line for the data $(x_1, y_1), ..., (x_n, y_n)$ must pass through $(\bar{x}, \bar{y})$. That is, show that $\bar{x}$ and $\bar{y}$ satisfy the linear equation $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1\bar{x}$.

Solution: (Added by Prajosh) We can begin from the equation

$$y = X\hat{\beta}$$

where

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} 1 & x \end{bmatrix}$$

The residual vector is given as
$\epsilon = y - X\hat{\beta}$
The residual vector ϵ is orthogonal to the column space of X and hence is orthogonal to 1. Then,

$$1.\epsilon = 1.(y - X\hat{\beta}) = 0$$

$$(y_1 + y_2 + ... + y_n) - \begin{bmatrix} n & \Sigma x \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \Sigma y - n\hat{\beta}_0 - \hat{\beta}_1 \Sigma x = n\bar{y} - n\hat{\beta}_0 - n\hat{\beta}_1\bar{x} = 0$$

$$\Rightarrow \bar{y} - \hat{\beta}_0 - \hat{\beta}_1\bar{x} = 0$$

$$\Rightarrow \bar{y} = \hat{\beta}_0 + \hat{\beta}_1\bar{x}$$

Hence proved.

9. Suppose $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ is an orthonormal basis of $\mathcal{V}$. Prove
(a) For any $\mathbf{u} \in \mathcal{V}$, we have $\mathbf{u} =< \mathbf{u}, \mathbf{e}_1 > \mathbf{e}_1 + < \mathbf{u}, \mathbf{e}_2 > \mathbf{e}_2 + ... + < \mathbf{u}, \mathbf{e}_n > \mathbf{e}_n$.
(b) $< a_1\mathbf{e}_1 + ... + a_n\mathbf{e}_n, b_1\mathbf{e}_1 + ... + b_n\mathbf{e}_n >= a_1b_1 + a_2b_2 + ...a_nb_n$.
(c) For any $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, we have $< \mathbf{u}, \mathbf{v} >=< \mathbf{u}, \mathbf{e}_1 >< \mathbf{v}, \mathbf{e}_1 > + ... + < \mathbf{u}, \mathbf{e}_n >< \mathbf{v}, \mathbf{e}_n >$

Solution: (Added by Prajosh)
(a) Suppose $u = k_1e_1 + k_2e_2 + ... + k_ne_n$. For any $i = 1, ..., n$, we have,

$$\begin{aligned} < \mathbf{u}, \mathbf{e}_i > &= \sum_{j=1, j\neq i}^{n} k_j < \mathbf{e}_j, \mathbf{e}_i > + k_i < \mathbf{e}_i, \mathbf{e}_i > \\ &= 0 + k_i(1) = k_i. \end{aligned}$$

Thus, $k_i =< \mathbf{u}, \mathbf{e}_i >$. This proves the desired result.
(b) Next, we have,

$$\left\langle \sum_{i=1}^{n} a_i\mathbf{e}_i, \sum_{j=1}^{n} b_j\mathbf{e}_j \right\rangle = \sum_{i,j=1}^{n} a_ib_j < \mathbf{e}_i, \mathbf{e}_j >= \sum_{i=1}^{n} a_ib_i < \mathbf{e}_i, \mathbf{e}_i > + \sum_{i\neq j} a_ib_j < \mathbf{e}_i, \mathbf{e}_j >$$

But $< \mathbf{e}_i, \mathbf{e}_j >= 0.$ for all $i \neq j$, and $< e_i, e_j >= 1$, if $i = j$. Hence proved the desired result.

(c) Here, we get,

$$< \mathbf{u}, \mathbf{v} > = \left\langle \sum_{i=1}^{n} < \mathbf{u}, \mathbf{e}_i > \mathbf{e}_i, \sum_{j=1}^{n} < \mathbf{v}, \mathbf{e}_j > \mathbf{e}_j \right\rangle \quad \text{[From part (a)]}$$
$$= \sum_{k=1}^{n} \Big(< \mathbf{u}, \mathbf{e}_k > . < \mathbf{v}, \mathbf{e}_k > \Big) \quad \text{[From part (b)].}$$

10. Consider the function $f : \mathbb{R} \to \mathbb{R}$. Sample input,output pairs of the function are given as $(1,1),(2,3),(3,1)$
    (a) Find the straight line $y = mx$, which has least mean squared error at the given sample points, using calculus.
    (b) Find the same line (find m), by writing a system of linear equations then projecting a vector onto the columnspace of system matrix. Find the vector that is being projected and find the vectorspace onto which that vector is projected.
    (c) Compare the data given in (a) and vector\vectorspace used for projection in (b). Convince yourself that the picture used in (a) and that in (b) are not same.

Solution: (Added by siva)

(a) The cost function to be minimized is the mean squared error.

$$E = \frac{1}{3}[(1-m)^2 + (3-2m)^2\,(1-3m)^2]$$
$$\frac{dE}{dm} = \frac{-1}{3}\left[2\,(1-m)) + 4\,(3-2m) + 6\,(1-3m)\right]$$

Since E is convex quadratic function, the mimimizer of E can be found by setting it's differentiation to zero.

$$\frac{dE}{dm} = 0$$
$$\implies \hat{m} = \frac{5}{7}$$

(b) The system of equations that we have from the data given, can be written using matrix notations as

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} m = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

Let $\mathbf{x}$ be $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ and $\mathbf{y}$ be $\begin{bmatrix} 1 & 3 & 1 \end{bmatrix}^T$.The least square solution of this system is $\hat{m} = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{y} = \frac{5}{7}$. Observe that $\mathbf{x}\hat{m} = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{y} = \hat{\mathbf{y}}$ is the projection of $\mathbf{y}$ onto the columnspace of $\mathbf{x}$. Hence, $\hat{m}$ is the coeffiecients that linearly combine the columns of $\mathbf{x}$ to produce the projected vector, $\hat{\mathbf{y}}$.

(c) The picture in (a) is $\mathbb{R}^2$-plane, which has the line $y = \frac{5}{7}x$. The picture used in (b) is the vectorspace $\mathbb{R}^3$, where we project the vector $\mathbf{y}$ onto the one-dimensional line defined as $\{c\mathbf{x}, c \in \mathbb{R}\}$.