

# Fundamentals of Linear Algebra and Optimization

Jean Gallier

## Homework 3

October 25, 2016

Francine Leech, Reffat Manzur, Chen Xiang

### Problem B1 (20 pts).

Let us prove this first for two subspaces.  $E = U_1 \oplus U_2$  if and only if

$$(1) \ E = U_1 + U_2$$

and

$$(2) \ U_1 \cap U_2 = \{0\}$$

For (1), every vector  $e \in E$  can be uniquely written as  $e = u_{11} + u_{21}$  with  $u_{11} \in U_1$  and  $u_{21} \in U_2$ .

For (2), let  $e \in U_1 \cap U_2$ . Since  $e \in U_1$  and  $e \in U_2$ , then we can write,

$$(1) \ e = e + 0 \text{ where } e \in U_1 \text{ and } 0 \in U_2 \text{ and}$$

$$(2) \ e = 0 + 0 \text{ where } 0 \in U_1 \text{ and } e \in U_2.$$

But  $e = u_{11} + u_{21}$  is unique, so  $e = 0$ . Since  $E = U_1 + U_2$ , we will check uniqueness. Suppose  $e = u_{11} + u_{21}$  and  $e = u_{12} + u_{22}$  where  $u_{11}, u_{12} \in U_1$  and  $u_{21}, u_{22} \in U_2$ . Then  $u_{11} + u_{21} = u_{12} + u_{22}$ , so  $u_{11} - u_{12} = u_{22} - u_{21}$ . Let  $x$  be a vector such that  $x = u_{11} - u_{12} = u_{22} - u_{21}$ . Then  $x \in U_1$  and  $x \in U_2$ , and  $u_{11} = u_{12}$  and  $u_{22} = u_{21}$ , so  $x \in U_1 \cap U_2 = \{0\}$ .

Now by induction we can extend our logic for any number of  $p \geq 2$  subspaces of some vector space  $E$ .

### Problem B2 (50 pts).

$$(1)$$

By definition an involution is a function  $f$  that is its own inverse.

$f : E \rightarrow E$  is in an involution when  $\forall x \in E : f(f(x)) = x$ . Since  $f^{-1} = f$ , we just have to check that  $f(f(x)) = x$  for all  $x$  in the domain of  $f$ .

Let us find the inverse of  $f(x) = a - x$  where  $a$  is any real number. Let,

$$f(x) = a - x$$

$$y = a - x$$

$$f^{-1} = a - y$$

$$x = a - y$$

$$y = a - x$$

so  $f(x) = f^{-1}(x)$ . So it's an inverse of itself.

(2)

$$\text{Let } u_1 = \frac{u+f(u)}{2} \text{ and } u_{-1} = \frac{u-f(u)}{2}.$$

$$\begin{aligned} \text{Next, we can find that we have something in both the spaces. } f(u_1) &= f\left(\frac{u+f(u)}{2}\right) = \\ \frac{f(u)+f(f(u))}{2} &= \frac{f(u)+u}{2} = u_1 \\ f(u_{-1}) &= f\left(\frac{u-f(u)}{2}\right) = \frac{f(u)-f(f(u))}{2} = \frac{f(u)-u}{2} = u_{-1}. \end{aligned}$$

Next we need to look at uniqueness. Let  $v_1 \in E_1$  and  $v_{-1} \in E_{-1}$ .

$$f(v_1) = v_1$$

$$f(v_1) = -v_1$$

$v_1 = -v_1$  can only have this if  $v_1 = -v_1 = 0$ .

$$(3) \text{ Let the basis of } E \text{ be } (\xi_i), i = 1, 2, \dots, n, \text{ where } \xi_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \text{ Define}$$

$$f(\xi_i) = \begin{cases} \xi_i & i = 1, 2, \dots, k \\ -\xi_i & i = k+1, k+2, \dots, n \end{cases}. \text{ Then } \forall u \in E, \text{ we can write as } u = \sum_i \lambda_i \xi_i \text{ and}$$

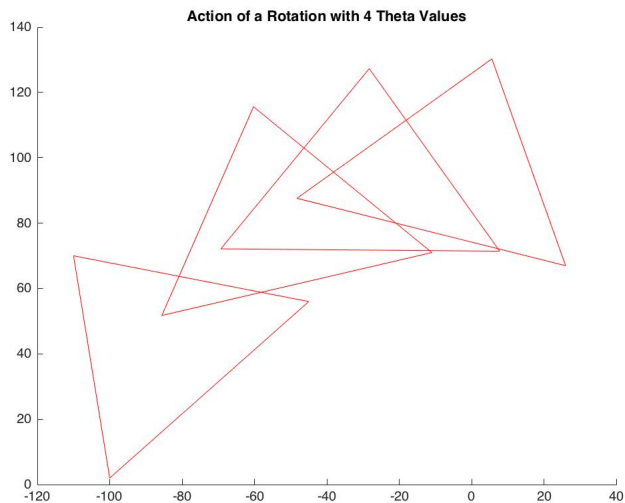
$$f(u) = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \\ -\xi_{k+1} \\ -\xi_{k+2} \\ \vdots \\ -\xi_n \end{pmatrix} = I_{k,n-k} \cdot u. \text{ And } (f \circ f)(u) = I_{k,n-k} I_{k,n-k} I_{k,n-k} u = I u = u. \text{ Thus } f \text{ is an}$$

involution.

Geometric interpretation of the action of  $f$  is that  $f(u)$  is the reflection of  $u$  across some axes. When  $k = n - 1$ , we can easily find that  $f(u)$  differs from  $u$  only in the last entry  $f(u)_i = -u_i$ , so it is the reflection across the  $n^{\text{th}}$  axis.

### Problem B3 (50 pts).

(1)



(2)

$$\begin{aligned} & \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{pmatrix} \cdot \begin{pmatrix} \cos(-1) & -\sin(-1) \\ \sin(-1) & \cos(-1) \end{pmatrix} \\ &= \begin{pmatrix} (\cos \theta \cdot \cos(-\theta)) \cdot (-\sin \theta \sin(-\theta)) & -\cos(\theta)0 \\ \sin(\theta)\cos(-\theta) + \cos\theta \sin(-\theta) & 1 \cos(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

(3) Because  $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$ ,  $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$ ,

we have

$$\begin{aligned} R_\alpha \circ R_\beta &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \cdot \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \\ &= R_\beta \circ R_\alpha = R_{\alpha+\beta} \end{aligned}$$

We can get

$$\begin{aligned} 1. & (R_\alpha \circ R_\beta) \circ R_\gamma = R_{\alpha+\beta+\gamma} = R_\alpha \circ (R_\beta \circ R_\gamma) \\ 2. & R_\alpha \circ I = I \circ R_\alpha = R_\alpha \\ 3. & R_\alpha \circ R_{-\alpha} = I \\ 4. & R_\alpha \circ R_\beta = R_{\alpha+\beta} = R_\beta \circ R_\alpha \end{aligned}$$

So rotations in the plane form a commutative group.

**Problem B4 (110 pts).**

(1) We can write  $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = R_{\theta, (a_1, a_2)} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$  as:

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

Then,

$$c_1 = \cos \theta c_1 - \sin \theta c_2 + a_1 \quad (1)$$

$$c_2 = \sin \theta c_1 + \cos \theta c_2 + a_2 \quad (2)$$

Using the result for  $(c_1, c_2)$  as given in the problem and by substituting, we will check if the equations are true for when  $\theta \neq k2\pi$ .

For  $c_1$ , we are given that

$$c_1 = \frac{1}{2 \sin \frac{\theta}{2}} (\cos(\frac{\pi - \theta}{2}) a_1 - \sin(\frac{\pi - \theta}{2}) a_2)$$

$\frac{1}{2} \frac{\cos(\frac{\pi - \theta}{2})}{\sin \frac{\theta}{2}}$  simplifies to  $\frac{1}{2}$ .

$\frac{-1}{2} \frac{\sin(\frac{\pi - \theta}{2})}{\sin \frac{\theta}{2}}$  simplifies to  $\frac{1}{2} \cot \frac{\theta}{2}$ .

so:

$$c_1 = \frac{1}{2} (a_1 - \cot \frac{\theta}{2} a_2)$$

Similarly, for  $c_2$ , it becomes

$$c_2 = \frac{1}{2} \cot \frac{\theta}{2} a_1 + \frac{1}{2} a_2$$

Now by substitution into equation (1), we get :

$$\begin{aligned}
& \frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2) \\
&= \cos \theta (\frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2)) - \sin \theta (\frac{1}{2} \cot \frac{\theta}{2} a_1 + \frac{1}{2} a_2) + a_1 \\
&= (\frac{\cos \theta}{2} - \frac{\sin \theta}{2} \cot \frac{\theta}{2} + 1) a_1 - \frac{1}{2} (\cos \theta \cot \frac{\theta}{2} + \sin \theta) a_2
\end{aligned}$$

Then,  $\cos \theta - \sin \theta \cot \frac{\theta}{2} = -1$  and  $\cos \theta \cot \frac{\theta}{2} + \sin \theta = \cot \frac{\theta}{2}$  so equation (1) on the right hand side simplifies to  $c_1 = \frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2)$  and matches the left hand side of equation (1).

For equation (2), we get:

$$\begin{aligned}
& \frac{1}{2} \cot \frac{\theta}{2} a_1 + \frac{1}{2} a_2 \\
&= \sin \theta (\frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2)) + \cos \theta (\frac{1}{2} \cot \frac{\theta}{2} a_1 + \frac{1}{2} a_2) + a_2 \\
&= (\frac{\sin \theta}{2} + (\frac{1}{2} \cos \theta \cot \frac{\theta}{2})) a_1 + ((\frac{-1}{2} \sin \theta \cot \frac{\theta}{2} + \frac{\cos \theta}{2} + 1) a_2
\end{aligned}$$

Thus, the right hand side of equation (2) simplifies and  $c_2 = \frac{1}{2} \cot \frac{\theta}{2} a_1 + \frac{1}{2} a_2$ .

It is then clear that the unique fixed point is  $(c_1, c_2)$  where  $\theta \neq k2\pi$  because  $\cot \frac{\theta}{2}$  is undefined for where  $\theta$  is a multiple of  $2\pi$ .

(2)

$$\begin{pmatrix} y'_1 + c_1 \\ y'_2 + c_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x'_1 + c_1 \\ x'_2 + c_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$y'_1 + \frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2) = \cos \theta x'_1 + \cos \theta (\frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2)) - \sin \theta x'_2 - \sin \theta (\frac{1}{2}(\cot \frac{\theta}{2} a_1 + a_2)) + a_1$$

$$= \cos \theta x'_1 - \sin \theta x'_2 + (\frac{\cos \theta}{2} + 1 - (\frac{\sin \theta \cot \frac{\theta}{2}}{2})) a_1 - \frac{1}{2} (\cos \theta \cot \frac{\theta}{2} + \sin \theta) a_2$$

Using  $\cos \theta - \sin \theta \cot \frac{\theta}{2} = -1$  and  $\cos \theta \cot \frac{\theta}{2} + \sin \theta = \cot \frac{\theta}{2}$ , the  $\frac{1}{2}(a_1 - \cot \frac{\theta}{2} a_2)$  cancel out from both sides of the equation and then

$$y'_1 = \cos \theta x'_1 - \sin \theta x'_2$$

$$y'_2 + c_2 = \sin \theta (x'_1 + c_1) + \cos \theta (x'_2 + c_2) + a_2$$

By substituting  $\frac{1}{2}(\cot\frac{\theta}{2}a_1 + a_2)$  for  $c_2$  and  $\frac{1}{2}(-\cot\frac{\theta}{2}a_2 + a_1)$  for  $c_1$ , we get

$$y'_2 = \sin\theta x'_1 + \cos\theta x'_2$$

Thus, the original affine map becomes

$$\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = R_\theta \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

(3) Use **Matlab** to show the action of the affine map  $R_{\theta,(a_1,a_2)}$  on a simple figure such as a triangle or a rectangle, for  $\theta = \pi/3$  and various values of  $(a_1, a_2)$ . Display the center  $(c_1, c_2)$  of the rotation.

What kind of transformations correspond to  $\theta = k2\pi$ , with  $k \in \mathbb{Z}$ ? The four centers of rotation in the figure are represented by \*.

When  $\theta = k2\pi$ , we have  $\cos\theta = 1$  and  $\sin\theta = 0$ , so  $y_1 = x_1 + a_1$  and  $y_2 = x_2 + a_2$  and this indicates the transformation is a translation.

(4)

$$R_{\theta,(a_1,a_2)}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & a_1 \\ \sin\theta & \cos\theta & a_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$$R_{-\theta,(b_1,b_2)}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) & b_1 \\ \sin(-\theta) & \cos(-\theta) & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) & b_1 \\ -\sin(\theta) & \cos(\theta) & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$$\begin{aligned} R_{\theta,(a_1,a_2)} \cdot R_{-\theta,(b_1,b_2)} &= \begin{pmatrix} \cos\theta & -\sin\theta & a_1 \\ \sin\theta & \cos\theta & a_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) & b_1 \\ -\sin(\theta) & \cos(\theta) & b_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta & \cos\theta b_1 - \sin\theta b_2 + a_1 \\ \cos\theta\sin\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta & \sin\theta b_1 + \cos\theta b_2 + a_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

So  $\cos \theta b_1 - \sin \theta b_2 + a_1 = 0$  and  $\sin \theta b_1 + \cos \theta b_2 + a_2 = 0$ .

Solve for  $b_1$ ,

$$\begin{aligned} b_1 &= \frac{-a_1 + \sin \theta b_2}{\cos \theta} \\ b_1 &= \frac{-a_1}{\cos \theta} + \frac{\sin \theta b_2}{\cos \theta} \\ b_1 &= -\sec \theta a_1 + \tan \theta b_2 \end{aligned}$$

Solve for  $b_2$ ,

$$\begin{aligned} b_2 &= \frac{-a_2 + \sin \theta b_1}{\cos \theta} \\ b_2 &= \frac{-a_2}{\cos \theta} - \frac{\sin \theta b_1}{\cos \theta} \\ b_2 &= \sec \theta a_2 - \tan \theta b_1 \end{aligned}$$

Now plug in  $b_2$  into  $b_1$ .

$$\begin{aligned} b_1 &= -\sec \theta a_1 + \tan \theta (\sec \theta a_2 - \tan \theta b_1) \\ b_1 &= \sec \theta a_1 - \tan \theta \sec \theta a_2 - \tan^2 \theta b_1 \\ (1 - \tan^2 \theta) b_1 &= \sec \theta a_1 - \tan \theta \sec \theta a_2 \\ b_1 &= \frac{-\tan \theta}{\sec \theta} a_2 - \cos \theta a_1 \\ b_1 &= -\sin \theta a_2 - \cos \theta a_1 \end{aligned}$$

Now plug  $b_1$  into  $b_2$  we get,  $b_2 = \sin \theta a_1 - \cos \theta a_1$ .

The final answer for  $b_1$  and  $b_2$  in terms of  $\theta$ ,  $a_1$ , and  $a_2$  is,

$$b_1 = -\sin \theta a_2 - \cos \theta a_1$$

$$b_2 = \sin \theta a_1 - \cos \theta a_1$$

(5)

iiiiiii Updated upstream  $R_{\beta, (b_1, b_2)}$  :

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \beta & -\sin \beta & b_1 \\ \sin \beta & \cos \beta & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$R_{\alpha, (a_1, a_2)}$  :

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & a_1 \\ \sin \alpha & \cos \alpha & a_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$$R_{\beta, (b_1, b_2)} \circ R_{\alpha, (a_1, a_2)} \rightarrow$$

$$\begin{pmatrix} \cos \beta & -\sin \beta & b_1 \\ \sin \beta & \cos \beta & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha & a_1 \\ \sin \alpha & \cos \alpha & a_2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\beta\cos\alpha - \sin\beta\sin\alpha & -\cos\beta\sin\alpha - \sin\beta\cos\alpha & \cos\beta a_1 - \sin\beta a_2 + b_1 \\ \sin\beta\cos\alpha + \cos\beta\sin\alpha & -\sin\beta\sin\alpha + \cos\beta\cos\alpha & \sin\beta a_1 + \cos\beta a_2 + b_2 \\ 0 & 0 & 1 \end{pmatrix}$$

Using the trig identities from part 4, we get:

$$\begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) & t_1 \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) & t_2 \\ 0 & 0 & 1 \end{pmatrix}$$

Indeed, this matrix represents  $R_{\alpha+\beta, (t_1, t_2)}$  where  $t_1 = \cos\beta a_1 - \sin\beta a_2 + b_1$  and  $t_2 = \sin\beta a_1 + \cos\beta a_2 + b_2$ .

Furthermore,

$$\begin{aligned} R_{\beta, (b_1, b_2)} \circ R_\alpha &\rightarrow \\ &\begin{pmatrix} \cos\beta & -\sin\beta & b_1 \\ \sin\beta & \cos\beta & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) & b_1 \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) & b_2 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

However,  $R_\alpha \circ R_{\beta, (b_1, b_2)} \rightarrow$

$$\begin{aligned} &\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta & b_1 \\ \sin\beta & \cos\beta & b_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) & \cos\alpha b_1 - \sin\alpha b_2 \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) & \sin\alpha b_1 + \cos\alpha b_2 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

so generally,

$$R_{\beta, (b_1, b_2)} \circ R_\alpha \neq R_\alpha \circ R_{\beta, (b_1, b_2)}$$

===== Use (4)-(5) to show that the affine maps of the plane defined in this problem form a nonabelian group denoted **SE**(2).

Prove that  $R_{\beta, (b_1, b_2)} \circ R_{\alpha, (a_1, a_2)}$  is not a translation (possibly the identity) iff  $\alpha + \beta \neq k2\pi$ , for all  $k \in \mathbb{Z}$ . Find its center of rotation when  $(a_1, a_2) = (0, 0)$ .

If  $\alpha + \beta = k2\pi$ , then  $R_{\beta, (b_1, b_2)} \circ R_{\alpha, (a_1, a_2)}$  is a pure translation. Find the translation vector of  $R_{\beta, (b_1, b_2)} \circ R_{\alpha, (a_1, a_2)}$ . Stashed changes

### Problem B5 (80 pts).

(1) Because  $U$  is a subspace of  $\mathbb{R}^n$ , then  $0 \in U$ , we have  $a = a + 0 \in \mathcal{A}$



Affine subspace of dimension 0 are points.

Affine subspace of dimension 1 are lines.

Affine subspace of dimension 2 are planes.

Any  $a + u \in \mathcal{A}$ ,  $a + v \in \mathcal{A}$ , then the affine combination is  $\lambda(a + u) + (1 - \lambda)(a + v) = a + \lambda u + (1 - \lambda)v$ , because  $U$  is subspace of  $\mathbb{R}^n$ ,  $\lambda u + (1 - \lambda)v \in U$ , then  $a + \lambda u + (1 - \lambda)v \in \mathcal{A} \Rightarrow \lambda(a + u) + (1 - \lambda)(a + v) \in \mathcal{A}$ . Affine subspace is closed under affine combinations.

(2) At least from (1) we know  $a \in \mathcal{A}$ ,  $\mathcal{A}$  is not empty. For any  $b \in \mathcal{A}$ , we can find  $u$ , write as  $b = a + u$ . Then  $b - a \in U$ ,  $a + U = a + b - a + U = b + U$ .

(3) Choose  $x - a \in U_a$ ,  $y - a \in U_a$ ,  $\lambda, \mu \in \mathbb{R}$ , then compute  $\lambda(x - a) + \mu(y - a) = \frac{1}{\lambda + \mu}(\frac{\lambda}{\lambda + \mu}(x - a) + \frac{\mu}{\lambda + \mu}(y - a)) = \frac{\lambda}{\lambda + \mu}x + \frac{\mu}{\lambda + \mu}y - a$ . Because  $\mathcal{A}$  is closed under affine combinations, we have  $\frac{\lambda}{\lambda + \mu}x + \frac{\mu}{\lambda + \mu}y \in \mathcal{A}$ , then  $\lambda(x - a) + \mu(y - a) = \frac{\lambda}{\lambda + \mu}x + \frac{\mu}{\lambda + \mu}y - a \in \mathbb{R}^n$ ,  $U_a$  is a subspace of  $\mathbb{R}^n$ .

For any  $u \in U_a$ , we can write as  $u = x - a = x + (b - a) - b$ . The sum of coefficients of  $x, b, a$  is  $1 - 1 + 1 = 1$ , and  $\mathcal{A}$  is closed under affine combinations, then  $u = x + (b - a) - b \in \mathcal{A} - b \Rightarrow u \in U_b$ .  $U_a$  does not depend on the choice of  $a \in \mathcal{A}$ ,  $U_a = U_b$  for all  $a, b \in \mathcal{A}$ . So  $U_a = U = \{y - x \in \mathbb{R}^n \mid x, y \in \mathcal{A}\}$  and  $\mathcal{A} = a + U$ , for any  $a \in \mathcal{A}$ .

(4) If  $\mathcal{A} \cap \mathcal{B} = (x)$ , then we can find  $x = a + u = b + v$  where  $a \in \mathcal{A}, b \in \mathcal{B}, u, v \in U$ . Write  $b = a + u - v$ , thus  $b \in \mathcal{A} \Rightarrow \mathcal{A} = \mathcal{B}$ , which is against  $\mathcal{A} \neq \mathcal{B}$ . We have  $\mathcal{A} \cap \mathcal{B} = \emptyset$  if  $\mathcal{A} \neq \mathcal{B}$  and  $\mathcal{A}, \mathcal{B}$  are parallel.

**Problem B6 (120 pts).** (Affine frames and affine maps)

(1)

We can write the  $\lambda$ 's and  $u$ 's as a unique linear combination such that,

$$\lambda_0 u_0 + \dots + \lambda_n u_n = 0$$

Since this is true, the only value for  $\sum_i \lambda_i$  is 0. We can rewrite this unique linear combination by subtracting  $u_0$  from each  $u_i$  to get

$$\lambda_1(u_1 - u_0) + \dots + \lambda_n(u_n - u_0) + \sum_i^n (u_i + u_0) = 0$$

Since  $\sum_i \lambda_i = 0$ , the linear combination of the difference between  $(u_i - u_0)$  is also linearly independent.

The subset of a linearly independent set of vectors is also a linearly independent set, so

$$\lambda_1(u_1 - u_0) + \dots + \lambda_m(u_m - u_0) + \sum_i^m (u_i + u_0) = 0$$

where  $m \leq n$  is linearly independent. Since the difference between the subset of vectors are linearly independent, then  $\sum_i \lambda_i = 0$  must be true. So  $u_0$  can be added to

$$\lambda_1(u_1 - u_0) + \dots + \lambda_m(u_m - u_0) + \sum_i^m (u_i + u_0) = 0$$

to get the linear combination,

$$\lambda_0 \hat{u}_0 + \lambda_1 \hat{u}_1 + \dots + \lambda_m \hat{u}_m = 0$$

Since  $\sum_i \lambda_i = 0$  is still true, then  $\hat{u}_0, \dots, \hat{u}_m$  are linearly independent.

(2)

We proved in 6.1 that the linear combination

$$\lambda_0 \hat{u}_0 + \lambda_1 \hat{u}_1 + \dots + \lambda_m \hat{u}_m = 0$$

is linearly independent and  $\sum_i \lambda_i = 0$ .

We can rewrite the linear combination by subtracting some  $u_i$  from each  $u$  to get

$$\lambda_0(\hat{u}_0 - \hat{u}_i) + \lambda_1(\hat{u}_1 - \hat{u}_i) + \dots + \lambda_m(\hat{u}_m - \hat{u}_i) = 0$$

The linear combination can be written by factoring out the  $u_i$  to get

$$\lambda_0 \hat{u}_0 + \lambda_1 \hat{u}_1 + \dots + \lambda_m \hat{u}_m - \sum_i \lambda_i \hat{u}_i = 0$$

The first set of terms of  $m - 1$  terms in the equations are a subset of the original  $m$  terms from where we started. Since a subset of a linearly independent set is linearly independent, then the  $m - 1$  terms are linearly independent and  $\sum_i^{m-1} \lambda_i = 0$ . This leaves the last term in the equation,  $\sum_i \lambda_i \hat{u}_i$ . We know that the  $u$ 's are unique, so they cannot equal 0, this means that the  $\lambda_i$ 's = 0. So, the  $m$  vectors  $u_j - u_i$  for  $j \in \{0, \dots, m\}$  with  $j - i \neq 0$  are linearly independent.

Any  $m + 1$  vectors  $(u_0, u_1, \dots, u_m)$  such that the  $m + 1$  vectors  $(\hat{u}_0, \dots, \hat{u}_m)$  are linearly independent are said to be *affinely independent*.

From (1) and (2), the vector  $(u_0, u_1, \dots, u_m)$  are affinely independent iff for any any choice of  $i$ , with  $0 \leq i \leq m$ , the  $m$  vectors  $u_j - u_i$  for  $j \in \{0, \dots, m\}$  with  $j - i \neq 0$  are linearly independent. If  $m = n$ , we say that  $n + 1$  affinely independent vectors  $(u_0, u_1, \dots, u_n)$  form an *affine frame* of  $\mathbb{R}^n$ .

iiiii Updated upstream (3) Because  $(u_0, u_1, \dots, u_n)$  is an affine frame of  $\mathbb{R}^n$ , then from above we can get a basis in  $\mathbb{R}^{n+1}$ , which is  $\hat{u}_0, \dots, \hat{u}_n$ . For any  $v \in \mathbb{R}^n$ , we denote  $\hat{v} = (v, 1)$ , then it can be uniquely written as the combination of  $(\hat{u}_i)$ , such that  $\hat{v} = \sum_i \lambda_i \hat{u}_i$ . From the last row of equation, we know there should be  $\sum_i \lambda_i = 1$ . Along with  $v = \sum_i \lambda_i u_i$ , we can

get the conclusion.

From  $e_i = u_i - u_0$ , for  $i = 1, \dots, n$ , substitute  $u_i$  with  $e_i$ , we get  $v = \sum_{i=1}^n \lambda_i(e_i + u_0) + \lambda_0 u_0 = \sum_i \lambda_i u_0 + \sum_{i=1}^n \lambda_i e_i = u_0 + \sum_{i=1}^n \lambda_i e_i$ . When  $u_i = u_0 + e_i$  for  $i = 1, \dots, n$ , denote the vector  $\hat{u}_i = (u_0 + e_i, 1)$ . Suppose  $\sum_i \lambda_i \hat{u}_i = 0$ , such that

$$\begin{aligned} \sum_i \lambda_i u_0 + \sum_i \lambda_i e_i &= 0 \\ \sum_i \lambda_i &= 0 \\ \Rightarrow \sum_i \lambda_i e_i &= 0 \\ \Rightarrow \lambda_i &= 0, \quad i = 1, 2, \dots, n \end{aligned}$$

So  $\hat{u}_0, \dots, \hat{u}_n$  are linearly independent, thus  $(u_0, u_1, \dots, u_n)$  is an affine frame of  $\mathbb{R}^n$ . From above, for any  $v \in \mathbb{R}^n$ , there is  $v = u_0 + x_1 e_1 + \dots + x_n e_n$ , substitute  $e_i$  with  $u_i$  get

$$v = (1 - (x_1 + \dots + x_n))u_0 + x_1 u_1 + \dots + x_n u_n,$$

===== (3) if  $(u_0, u_1, \dots, u_n)$  is an affine frame of  $\mathbb{R}^n$ , then prove that for every vector  $v \in \mathbb{R}^n$ , there is a unique  $(n+1)$ -tuple  $(\lambda_0, \lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n+1}$ , with  $\lambda_0 + \lambda_1 + \dots + \lambda_n = 1$ , such that

$$v = \lambda_0 u_0 + \lambda_1 u_1 + \dots + \lambda_n u_n.$$

The scalars  $(\lambda_0, \lambda_1, \dots, \lambda_n)$  are called the *barycentric* (or *affine*) *coordinates* of  $v$  w.r.t. the affine frame  $(u_0, u_1, \dots, u_n)$ .

(3.1)

If we write  $e_i = u_i - u_0$ , for  $i = 1, \dots, n$ , then prove that we have

$$v = u_0 + \lambda_1 e_1 + \dots + \lambda_n e_n,$$

and since  $(e_1, \dots, e_n)$  is a basis of  $\mathbb{R}^n$  (by (1) & (2)), the  $n$ -tuple  $(\lambda_1, \dots, \lambda_n)$  consists of the standard coordinates of  $v - u_0$  over the basis  $(e_1, \dots, e_n)$ .

(3.2)

From above, where

$$\hat{v} = \lambda_0 \hat{u}_0 + \lambda_1 \hat{u}_1 + \dots + \lambda_n \hat{u}_n$$

are linearly independent. We can rewrite above as,

$$\hat{v} = \lambda_1(\hat{u}_1 - \hat{u}_0) + \lambda_2(\hat{u}_2 - \hat{u}_0) + \dots + \lambda_n(\hat{u}_n - \hat{u}_0) + \sum_i^n \lambda_i \hat{u}_0$$

We can replace each  $u_i - u_0$  with  $e_i$  such that  $e_i = u_i - u_0$ . Now,

$$\hat{v} = \lambda_1(e_1) + \lambda_2(e_2) + \dots + \lambda_n(e_n) + \sum_i^n \lambda_i u_0$$

From 6.1, we know that the subset of a linearly independent set is linearly independent. Thus the first  $n - 1$  is linearly independent, so the  $\sum_{i=1}^{n-1} \lambda_i = 0$ . The last  $n$ th term in the equation,  $\sum_{i=1}^n \lambda_i u_0$  must equal 0. Since the  $u_0$  is unique, the  $\lambda_i$ 's must equal 0.

Conversely, for any vector  $u_0 \in \mathbb{R}^n$  and for any basis  $(e_1, \dots, e_n)$  of  $\mathbb{R}^n$ , let  $u_i = u_0 + e_i$  for  $i = 1, \dots, n$ . Prove that  $(u_0, u_1, \dots, u_n)$  is an affine frame of  $\mathbb{R}^n$ , and for any  $v \in \mathbb{R}^n$ , if

$$v = u_0 + x_1 e_1 + \dots + x_n e_n,$$

with  $(x_1, \dots, x_n) \in \mathbb{R}^n$  (unique), then

$$v = (1 - (x_1 + \dots + x_n))u_0 + x_1 u_1 + \dots + x_n u_n,$$

so that  $(1 - (x_1 + \dots + x_n), x_1, \dots, x_n)$ , are the barycentric coordinates of  $v$  w.r.t. the affine frame  $(u_0, u_1, \dots, u_n)$ .

(3.3)

The above shows that there is a one-to-one correspondence between affine frames  $(u_0, \dots, u_n)$  and pairs  $(u_0, (e_1, \dots, e_n))$ , with  $(e_1, \dots, e_n)$  a basis. Given an affine frame  $(u_0, \dots, u_n)$ , we obtain the basis  $(e_1, \dots, e_n)$  with  $e_i = u_i - u_0$ , for  $i = 1, \dots, n$ ; given the pair  $(u_0, (e_1, \dots, e_n))$  where  $(e_1, \dots, e_n)$  is a basis, we obtain the affine frame  $(u_0, \dots, u_n)$ , with  $u_i = u_0 + e_i$ , for  $i = 1, \dots, n$ . There is also a one-to-one correspondence between barycentric coordinates w.r.t. the affine frame  $(u_0, \dots, u_n)$  and standard coordinates w.r.t. the basis  $(e_1, \dots, e_n)$ . The barycentric coordinates  $(\lambda_0, \lambda_1, \dots, \lambda_n)$  of  $v$  (with  $\lambda_0 + \lambda_1 + \dots + \lambda_n = 1$ ) yield the standard coordinates  $(\lambda_1, \dots, \lambda_n)$  of  $v - u_0$ ; the standard coordinates  $(x_1, \dots, x_n)$  of  $v - u_0$  yield the barycentric coordinates  $(1 - (x_1 + \dots + x_n), x_1, \dots, x_n)$  of  $v$ .

~~~~~ Stashed changes (4)

We can write  $f(u) = f(u_0 \lambda_0 + u_1 \lambda_1 + \dots + u_n \lambda_n) = \lambda_0 f(u_0) + \lambda_1 f(u_1) + \dots + \lambda_n f(u_n) = \lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n$ . So we can define a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  as above, such that  $f(\lambda_0 u_0 + \lambda_1 u_1 + \dots + \lambda_n u_n) = \lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n$ . The function is unique, however we have to check if it is affine.

To check if it is affine, we first pick  $w_1, \dots, w_p \in \mathbb{R}^n$  and a set of  $\mu_1 + \dots + \mu_p = 1$ . We can write linear combination of  $\mu$  and  $w$ , such that,  $f(\mu_1 w_1 + \dots + \mu_p w_p) = \mu_1 f(w_1) + \dots + \mu_p f(w_p)$ . Now we can rewrite all the  $w$ 's as the affine frame, such that,

$$w_1 = \sum_{i=0}^n \lambda_i^1 \mu_i, \dots, w_p = \sum_{i=0}^n \lambda_i^p \mu_i$$

Now we can expand and regroup in terms of the frame vectors.

$$f(u_1(\sum_{i=0}^n \lambda_i^1 \mu_i)) + \dots + \mu_p(\sum_{i=0}^n \lambda_i^p \mu_i) = f((\sum_{j=1}^p \mu_j \lambda_0^j) u_0, \dots, (\sum_{j=1}^p \mu_j \lambda_n^j) u_n)$$

Now we take the coefficient out and set equal to 1, this works out because  $\sum_i \lambda_i = 1$  and  $\sum_i \mu_i = 1$ .

$$1 = \left(\sum_{j=1}^p \mu_j \lambda_0^j\right) v_0 + \dots + \left(\sum_{j=1}^p \mu_j \lambda_n^j\right) v_n$$

Now regroup,

$$1 = \mu_1 \left(\sum_{i=1}^n \lambda_i^1 v_i\right) + \dots + \mu_p \left(\sum_{i=1}^n \lambda_i^p v_i\right)$$

Because of our original definition,

$$w_1 = \sum_{i=0}^n \lambda_i^0 \mu_i, \dots, w_p = \sum_{i=0}^n \lambda_i^p \mu_i$$

we can conclude that

$$1 = \mu_1 f(w_1) + \dots + \mu_p f(w_p)$$

so  $f$  is an affine combination.

(5) Let  $(a_0, \dots, a_n)$  be any affine frame in  $\mathbb{R}^n$  and let  $(b_0, \dots, b_n)$  be any  $n+1$  points in  $\mathbb{R}^n$ . From (4), we know there is a unique affine map  $f$  such that  $f(a_i) = b_i$ ,  $i = 0, \dots, n$ . From (3), we know for any  $b_i$ , we can write as the affine combination of  $(a_i)$ , such that  $b_i = \sum_j \lambda_j^i a_j$ ,  $\sum_j \lambda_j^i = 1$ . Thus we can write those equations as

$$\begin{aligned} \begin{pmatrix} b_0 & b_1 & \dots & b_n \\ 1 & 1 & \dots & 1 \end{pmatrix} &= \begin{pmatrix} a_0 & a_1 & \dots & a_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} \lambda_0^0 & \lambda_0^1 & \dots & \lambda_0^n \\ \lambda_1^0 & \lambda_1^1 & \dots & \lambda_1^n \\ \lambda_2^0 & \lambda_2^1 & \dots & \lambda_2^n \\ \vdots & \dots & \dots & \vdots \\ \lambda_n^0 & \lambda_n^1 & \dots & \lambda_n^n \end{pmatrix} \\ \Rightarrow A &= (\hat{a}_0 \quad \hat{a}_1 \quad \dots \quad \hat{a}_n)^{-1} \cdot (\hat{b}_0 \quad \hat{b}_1 \quad \dots \quad \hat{b}_n) \end{aligned}$$

the  $(n+1) \times (n+1)$  matrix  $A$  corresponding to the unique affine map  $f$  such that

$$f(a_i) = b_i, \quad i = 0, \dots, n,$$

is given by

$$A = \begin{pmatrix} \hat{b}_0 & \hat{b}_1 & \dots & \hat{b}_n \end{pmatrix} (\hat{a}_0 \quad \hat{a}_1 \quad \dots \quad \hat{a}_n)^{-1}.$$

When  $(a_0, \dots, a_n)$  is the canonical affine frame with  $a_i = e_{i+1}$  for  $i = 0, \dots, n-1$  and  $a_n = (0, \dots, 0)$  there is

$$(\hat{a}_0 \quad \hat{a}_1 \quad \dots \quad \hat{a}_n) = \begin{pmatrix} a_0 & a_1 & \dots & a_n \\ 1 & 1 & \dots & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix}$$

$$\begin{aligned}
(\hat{a}_0 \quad \hat{a}_1 \quad \cdots \quad \hat{a}_n) \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix} \\
&= I \\
\Rightarrow (\hat{a}_0 \quad \hat{a}_1 \quad \cdots \quad \hat{a}_n)^{-1} &= \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix}.
\end{aligned}$$

(6) Choose  $u_i \in f(\mathcal{A})$ ,  $u_i = f(a_i)$ ,  $a_i \in \mathcal{A}$  and  $\lambda_i \in \mathbb{R}$ ,  $\sum_i \lambda_i = 1$ . The affine combination of  $u_i$  is  $\sum_i \lambda_i u_i = f(\sum_i \lambda_i a_i)$ . Because  $\mathcal{A}$  is affine subspace then  $\sum_i \lambda_i a_i \in \mathcal{A}$ , which means  $\sum_i \lambda_i u_i \in f(\mathcal{A})$ , then  $f(\mathcal{A})$  is affine subspace too. To do the same thing,  $v_i \in f^{-1}(\mathcal{B})$ ,  $v_i = f^{-1}(b_i)$ ,  $b_i \in \mathcal{B}$  and  $\mu_i \in \mathbb{R}$ ,  $\sum_i \mu_i = 1$ . The affine combination of  $v_i$  is  $\sum_i \mu_i v_i = f^{-1}(\sum_i \mu_i b_i)$ . Because  $\mathcal{B}$  is affine subspace then  $\sum_i \mu_i b_i \in \mathcal{B}$ , which means  $\sum_i \mu_i v_i \in f^{-1}(\mathcal{B})$ , then  $f^{-1}(\mathcal{B})$  is affine subspace too.

**Problem B7 (30 pts).** Let  $A$  be any  $n \times k$  matrix

(1) If choose  $u \in \ker(A)$ ,  $Au = 0$  then

$$\begin{aligned}
&\Rightarrow (A^\top A)u = A^\top(Au) = 0 \\
&\Rightarrow u \in \ker(A^\top A) \\
&\Rightarrow \ker(A) \subseteq \ker(A^\top A)
\end{aligned}$$

For the opposite, choose  $v \in \ker(A^\top A)$  then

$$\begin{aligned}
&\Rightarrow (A^\top A)v = A^\top(Av) = 0 \\
&\Rightarrow Av \in \ker(A^\top)
\end{aligned}$$

We need to prove  $u \in \ker(A)$ , if not suppose  $Au = x \neq 0$ , then  $(A^\top x)^\top = x^\top A = 0$ , multiply  $u$  on both sides, we have  $x^\top Au = x^\top x = 0 \Rightarrow x = 0$ , which is against the assumption, thus  $u \in \ker(A) \Rightarrow \ker(A^\top A) \subseteq \ker(A)$ . From  $\ker(A) \subseteq \ker(A^\top A)$ ,  $\ker(A^\top A) \subseteq \ker(A)$ , get  $\ker(A^\top A) = \ker(A)$ . From the equation  $\dim(\ker(A))\text{rank}(A) = \dim(E) = k$ , we can easily find  $\text{rank}(A^\top A) = \text{rank}(A) = k - \dim(\ker(A)) = k - \dim(\ker(A^\top A))$ . We can use the same way to prove  $\ker(AA^\top) = \ker(A^\top)$  and  $\text{rank}(AA^\top) = \text{rank}(A^\top)$ .

(2) From above, we know  $\text{rank}(A^\top A) = \text{rank}(A)$ .  $\text{rank}(A) \leq \min\{k, n\} = k$ , because  $A$  has independent column vectors  $(a_i)$   $i = 1, 2, \dots, k$ , then  $\text{rank}(A) = k \Rightarrow \text{rank}(A^\top A) = k$ .  $A^\top A$  has full rank, it is invertible.

First we need to prove two things:

$$\begin{aligned}(AB)^\top &= B^\top A^\top \\ (A^{-1})^\top &= (A^\top)^{-1}\end{aligned}$$

Here is the simple proof:

$$\begin{aligned}(AB)_{ij}^\top &= \sum_k a_{jk} b_{ki} = \sum_k b_{ki} a_{jk} = (B^\top A^\top)_{ij} \\ (A^{-1})^\top A^\top &= (AA^{-1})^\top = I \\ \Rightarrow (A^{-1})^\top &= (A^\top)^{-1}\end{aligned}$$

Then

$$\begin{aligned}P^2 &= A(A^\top A)^{-1} A^\top \cdot A(A^\top A)^{-1} A^\top \\ &= A(A^\top A)^{-1} (A^\top A) (A^\top A)^{-1} \\ &= A(A^\top A)^{-1} A^\top \\ &= P \\ P^\top &= (A(A^\top A)^{-1} A^\top)^\top \\ &= A(A^\top A)^{-1} A^\top \\ &= A((A^\top A)^\top)^{-1} A^\top \\ &= A(A^\top A)^{-1} A^\top \\ &= P\end{aligned}$$

When  $k = 1$ ,  $P$  is symmetric matrix with  $\text{trace}(P) = 1$ . If  $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$  then  $P$  looks like

$$P = \frac{1}{\sum_i a_i^2} \begin{pmatrix} a_1^2 & a_1 a_2 & a_1 a_3 & \cdots & a_1 a_n \\ a_1 a_2 & a_2^2 & a_2 a_3 & \cdots & a_2 a_n \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ a_1 a_n & a_2 a_n & a_3 a_n & \cdots & a_n^2 \end{pmatrix}.$$

(3) Choose any  $x \in \mathbb{R}^n$ , the image of  $P$  is  $Px = A(A^\top A)^{-1} A^\top x = A[(A^\top A)^{-1} A^\top x]$ , there is  $y \in \mathbb{R}^k$  makes  $Px = Ay$ , so the image of  $P$  is the subspace  $V$  spanned by  $a_1, a_2, \dots, a_n$ . When  $u \in U$ ,  $u \in \ker(P)$ , we can get  $Pu = A(A^\top A)^{-1} A^\top u = 0$ , multiply  $A^\top$  on both sides,  $A^\top A(A^\top A)^{-1} A^\top u = A^\top u = 0 \Rightarrow \ker(P) \subseteq \ker(A^\top)$ . So the nullspace  $U$  of  $P$  is the set of vectors such that  $A^\top u = 0$ . Geometric interpretation of  $U$  is that  $U$  contains vectors that are orthogonal to the image of  $A$ .

First we need to prove that for any  $x \in \mathbb{R}^n$ ,  $Px$  is the closest vector in  $V$ . Because  $(Px -$

$x)^\top(Px) = (x^\top P^\top - x^\top)Px = 0$ , so  $Px - x$  is perpendicular to  $Px$ , which means  $P$  is a projection of  $\mathbb{R}^n$  onto subspace spanned by  $(a_i) \ i = 1, 2, \dots, k$ .

Then what we only need to do is to prove  $V^0 = U$ . Choose  $u \in U$ , from above we know  $A^\top u = 0$ , thus any  $Ax \in V$  we have  $(Ax)^\top y = x^\top A^\top y = 0 \Rightarrow U \subseteq V^0$ . For any  $v \in V^0$ , we have  $(Ax)^\top v = x^\top A^\top v = 0$ , because we can choose any  $x \in \mathbb{R}^n$ , then must have the result  $A^\top v = 0 \Rightarrow V^0 \subseteq U$ . (Otherwise we can construct  $(e_i) \ i = 1, 2, \dots, n$  and then  $(e_1, e_2, \dots, e_n)^\top A^\top v = I \cdot A^\top v = 0$ ). As a conclusion  $U = V^0 \Rightarrow \mathbb{R}^n = V^0 \oplus V = U \oplus V$ .

**TOTAL: 460 points.**

## APPENDIX

### Problem 3.3 Code

```

1 %Theta = pi/6
2 theta = pi/6;
3 R = [cos(theta) -sin(theta);
4      sin(theta) cos(theta)];
5
6 x = [2;70;56];
7 y = [100;110;45];
8 xy = [x y];
9 Rxy = R*xy';
10 Rxy = Rxy';
11 x = Rxy(:,1);
12 y = Rxy(:,2);
13 line([x(1),x(2)], [y(1), y(2)], 'Color', 'r');
14 hold on;
15 line([x(2),x(3)], [y(2), y(3)], 'Color', 'r');
16 line([x(3),x(1)], [y(3), y(1)], 'Color', 'r');
17 hold off;
18
19 %Theta = pi/4
20 theta = pi/4;
21 R = [cos(theta) -sin(theta);
22      sin(theta) cos(theta)];
23
24 x = [2;70;56];
25 y = [100;110;45];
26 xy = [x y];
27 Rxy = R*xy';
28 Rxy = Rxy';
29 x = Rxy(:,1);

```



```

30 y = Rxy(:,2);
31 line([x(1),x(2)], [y(1), y(2)], 'Color', 'r');
32 hold on;
33 line([x(2),x(3)], [y(2), y(3)], 'Color', 'r');
34 line([x(3),x(1)], [y(3), y(1)], 'Color', 'r');
35 hold off;
36
37 %Theta = pi/3
38 theta = pi/3;
39 R = [cos(theta) -sin(theta);
40      sin(theta) cos(theta)];
41
42 x = [2;70;56];
43 y = [100;110;45];
44 xy = [x y];
45 Rxy = R*xy';
46 Rxy = Rxy';
47 x = Rxy(:,1);
48 y = Rxy(:,2);
49 line([x(1),x(2)], [y(1), y(2)], 'Color', 'r');
50 hold on;
51 line([x(2),x(3)], [y(2), y(3)], 'Color', 'r');
52 line([x(3),x(1)], [y(3), y(1)], 'Color', 'r');
53 hold off;
54
55 %Theta = pi/2
56 theta = pi/2;
57 R = [cos(theta) -sin(theta);
58      sin(theta) cos(theta)];
59
60 x = [2;70;56];
61 y = [100;110;45];
62 xy = [x y];
63 Rxy = R*xy';
64 Rxy = Rxy';
65 x = Rxy(:,1);
66 y = Rxy(:,2);
67 line([x(1),x(2)], [y(1), y(2)], 'Color', 'r');
68 hold on;
69 line([x(2),x(3)], [y(2), y(3)], 'Color', 'r');
70 line([x(3),x(1)], [y(3), y(1)], 'Color', 'r');
71 hold off;

```