$\boxed{\textbf{Fall 2016} \quad \textbf{CIS 515}}$

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) $e = e + 0$ where $e \in U_1$ and $0 \in U_2$ and

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

But $e = u_{11} + u_{21}$ is unique, so $e = 0$. Since $E = U_1 + U_P$, 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 $u_{11} + u_{21} = u_{12} + u22$, 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.

$f : E \to E$ 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 $f(x) = b - x$ where b 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)

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

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

Next we need to look at uniqueness. Let $v_1 \in E_1 and \in E_{-1}$.
$f(v_1) = v_1$
$f(v_1) = -v_1$
$v_1 = -v_1$ can only have this if $v_1 = -v_1 = 0$.

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

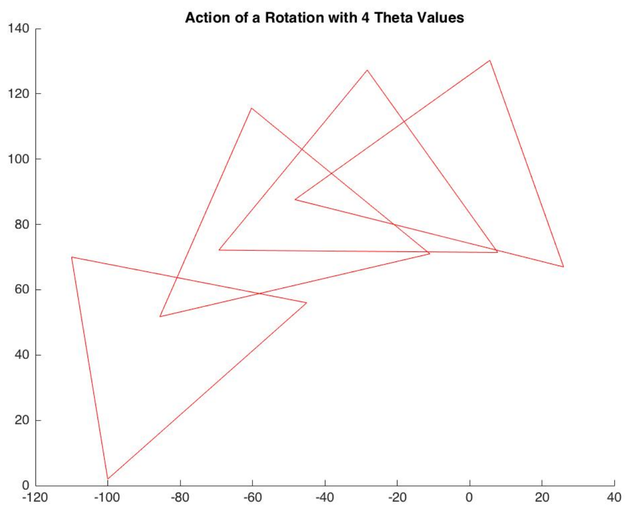
$f(\xi_i) = \begin{cases} \xi_i \; i = 1, 2, \cdots, k \\ -\xi_i \; i = k+1, k+2, \cdots, n \end{cases}$. Then $\forall u \in E$, we can write as $u = \sum_i \lambda_i \xi_i$ and

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

Geometric interpretation of the action of f is that $f(u)$ is the reflection of u across some axises. 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^{th} axis.

Problem B3 (50 pts).

(1)



(2)

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix}$$

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

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

(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,

$$c_1 = cos\theta c_1 - sin\theta c_2 + a_1 \tag{1}$$

$$c_2 = sin\theta c_1 + cos\theta c_2 + a_2 \tag{2}$$

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

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

$\frac{1}{2}\frac{cos(\frac{\pi-\theta}{2})}{sin\frac{\theta}{2}}$ 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 :

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

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

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

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:

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

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

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

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 θ is a multiple of 2π.

(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 $\frac{1}{2}\left(cot\frac{\theta}{2}a_1 + a_2\right)$ for c_2 and $\frac{1}{2}\left(-cot\frac{\theta}{2}a_2 + a_1\right)$ 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)

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

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

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

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

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

$$= \begin{pmatrix} \cos^2\theta + \sin^2\theta & cos\theta\sin\theta - cos\theta\sin\theta & \cos\theta b_1 - \sin\theta b_2 + a_1 \\ cos\theta\sin\theta - cos\theta\sin\theta & \sin^2\theta + \cos^2\theta & \sin\theta b_1 + \cos\theta b_2 + a_2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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,
$b_1 = \frac{-a_1 + \sin\theta b_2}{\cos\theta}$
$b_1 = \frac{-a_1}{\cos\theta} + \frac{\sin\theta b_2}{\cos\theta}$
$b_1 = -\sec\theta a_1 + \tan\theta b_2$

Solve for b_2,
$b_2 = \frac{-a_2 + \sin\theta b_1}{\cos\theta}$
$b_2 = \frac{-a_2}{\cos\theta} - \frac{\sin\theta b_1}{\cos\theta}$
$b_2 = \sec\theta a_2 - \tan\theta b_1$

Now plug in b_2 into b_1.
$b_1 = -\sec\theta a_1 + \tan\theta(\sec\theta a_2 - \tan\theta b_1)$
$b_1 = \sec\theta a_1 - \tan\theta\sec\theta a_2 - \tan^2\theta b_1$
$(1 - tan^2\theta)b_1 = \sec\theta a_1 - \tan\theta\sec\theta a_2$
$b_1 = \frac{-\tan\theta}{\sec\theta}a_2 - \cos\theta a_1$
$b_1 = -\sin\theta a_2 - \cos\theta a_1$
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 θ, 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)
$R_{\beta,(b_1,b_2)}$:

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

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

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

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

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

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

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,

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

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

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

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

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

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

so generally,

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

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 λ's and u's as a unique linear combination such that,

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

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

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

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

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

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

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

to get the linear combination,

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

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

(2)

We proved in 6.1 that the linear combination

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

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

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

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

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

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 λ_i's $= 0$. So, the m vectors $u_j - u_i$ for $j \in \{0, \ldots, m\}$ with $j - i \neq 0$ are linearly independent.

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

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

(3) Because $(u_0, u_1, \ldots, u_n)$ is an affine frame of $\mathbb{R}^n$, then from above we can get a basis in $\mathbb{R}^{n+1}$, which is $\widehat{u}_0, \ldots, \widehat{u}_n$. For any $v \in \mathbb{R}^n$, we denote $\widehat{v} = (v, 1)$, then it can be uniquely written as the combination of $(\widehat{u}_i)$, such that $\widehat{v} = \sum_i \lambda_i\widehat{u}_i$. 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 $i = 1, \ldots, n$, 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 $i = 1, \ldots, n$, denote the vector

$\widehat{u_i} = (u_0 + e_i, 1)$. Suppose $\sum_i \lambda_i \widehat{u_i} = 0$, such that

$$\sum_i \lambda_i u_0 + \sum_i \lambda_i e_i = 0$$
$$\sum_i \lambda_i = 0$$
$$\Rightarrow \sum_i \lambda_i e_i = 0$$
$$\Rightarrow \lambda_i = 0,\ i = 1, 2, \cdots, n$$

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

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

(4)
We can write $f(u) = f(u_0\lambda_0 + u_1\lambda_1 + \ldots + u_n\lambda_n) = \lambda_0 f(u_0) + \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) = \lambda_0 v_0 + \lambda_1 v_1 + \ldots + \lambda_n v_n$. So we can define a function $f : \mathbb{R}^n \to \mathbb{R}^m$ as above, such that $f(\lambda_0 u_0 + \lambda_1 u_1 + \ldots + \lambda_n u_n) = \lambda_0 v_0 + \lambda_1 v_1 + \ldots + \lambda_n v_n$. The function is unique, however we have to check if it is affine.

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

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

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

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

Now we take the coefficient out and set equal to 1, this works out because $\sum_i \lambda_1 = 1$ and $\sum_i \mu_1 = 1$.

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

Now regroup,

$$1 = \mu_1(\sum_{i=1}^{n} \lambda_i^1 v_i) + \ldots + \mu_p(\sum_{i=1}^{n} \lambda_i^n v_i)$$

Because of our original definition,

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

we can conclude that

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

so f is an affine combination.

(5) Let $(a_0, \ldots, a_n)$ be any affine frame in $\mathbb{R}^n$ and let $(b_0, \ldots, b_n)$ be any $n+1$ points in $\mathbb{R}^n$. From (4), we know there is a unique affine map f such that $f(a_i) = b_i,\ i = 0, \ldots, n,$. 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

$$\begin{pmatrix} b_0 & b_1 & \cdots & b_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \lambda_0^0 & \lambda_0^1 & \cdots & \lambda_0^n \\ \lambda_1^0 & \lambda_1^1 & \cdots & \lambda_1^n \\ \lambda_2^0 & \lambda_2^1 & \cdots & \lambda_2^n \\ \vdots & \cdots & \cdots & \vdots \\ \lambda_n^0 & \lambda_n^1 & \cdots & \lambda_n^n \end{pmatrix}$$

$$\Rightarrow A = \begin{pmatrix} \widehat{a}_0 & \widehat{a}_1 & \cdots & \widehat{a}_n \end{pmatrix}^{-1} . \begin{pmatrix} \widehat{b}_0 & \widehat{b}_1 & \cdots & \widehat{b}_n \end{pmatrix}$$

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

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

is given by

$$A = \begin{pmatrix} \widehat{b}_0 & \widehat{b}_1 & \cdots & \widehat{b}_n \end{pmatrix} \begin{pmatrix} \widehat{a}_0 & \widehat{a}_1 & \cdots & \widehat{a}_n \end{pmatrix}^{-1}.$$

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

$$\begin{pmatrix} \widehat{a}_0 & \widehat{a}_1 & \cdots & \widehat{a}_n \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \widehat{a}_0 & \widehat{a}_1 & \cdots & \widehat{a}_n \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix}$$

$$= I$$

$$\Rightarrow \begin{pmatrix} \widehat{a}_0 & \widehat{a}_1 & \cdots & \widehat{a}_n \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix}.$$

(6) Choose $u_i \in f(\mathcal{A}),\ u_i = f(a_i),\ a_i\imath\mathcal{A}$ 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 $\sum_i \mu_i v_i = f(\sum_i \mu_i b_i)$.Because $\mathcal{B}$ is affine subspace then $\sum_i \mu_i b_i \in \mathcal{B}$, which means$\sum_i \mu_i v_i \in f(\mathcal{B})$, 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

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

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 $\text{rank}(AA^\top) = \mathrm{A}^\top$.

(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) $i = 1, 2, \cdots k$, 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:

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

Then

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

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

$P = \frac{1}{\sum_i a_i^2} \begin{pmatrix} a_1^2 & a_1a_2 & a_1a_3 & \cdots & a_1a_n \\ a_1a_2 & a_2^2 & a_2a_3 & \cdots & a_2a_n \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ a_1a_n & a_2a_n & a_3a_n & \cdots & a_n^2 \end{pmatrix}$.

(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 $a_1, a_2, \cdots, a_n$. 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) $i = 1, 2, \cdots k$.
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 $A^\top x = 0 \Rightarrow V^0 \subseteq U$. (Otherwise we can construct (e_i) $i = 1, 2, \cdots n$ and then $(e_1, e_2, \cdots e_n)^\top A^\top v = I \cdot A^\top v = 0$). As a conclusion $U = V^0 \Rightarrow \mathbb{R}^n = V^0 \oplus V = U \oplus V$.

TOTAL: 460 points.