

# Fundamentals of Linear Algebra and Optimization

Jean Gallier

## Homework 3

October 2, 2015; Due October 20, 2015

Beginning of class

**Problem B1 (20 pts).** Let  $U_1, \dots, U_p$  be any  $p \geq 2$  subspaces of some vector space  $E$ . Prove that  $U_1 + \dots + U_p$  is a direct sum iff

$$U_i \cap \left( \sum_{j=1}^{i-1} U_j \right) = (0), \quad i = 2, \dots, p.$$

**Problem B2 (50 pts).** Given any vector space  $E$ , a linear map  $f: E \rightarrow E$  is an *involution* if  $f \circ f = \text{id}$ .

- (1) Prove that an involution  $f$  is invertible. What is its inverse?
- (2) Let  $E_1$  and  $E_{-1}$  be the subspaces of  $E$  defined as follows:

$$\begin{aligned} E_1 &= \{u \in E \mid f(u) = u\} \\ E_{-1} &= \{u \in E \mid f(u) = -u\}. \end{aligned}$$

Prove that we have a direct sum

$$E = E_1 \oplus E_{-1}.$$

*Hint.* For every  $u \in E$ , write

$$u = \frac{u + f(u)}{2} + \frac{u - f(u)}{2}.$$

(3) If  $E$  is finite-dimensional and  $f$  is an involution, prove that there is some basis of  $E$  over which the matrix of  $f$  is of the form

$$I_{k,n-k} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix},$$

where  $I_k$  is the  $k \times k$  identity matrix (similarly for  $I_{n-k}$ ) and  $k = \dim(E_1)$ . Can you give a geometric interpretation of the action of  $f$  (especially when  $k = n - 1$ )?

**Problem B3 (50 pts).** A rotation  $R_\theta$  in the plane  $\mathbb{R}^2$  is given by the matrix

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

(1) Use `Matlab` to show the action of a rotation  $R_\theta$  on a simple figure such as a triangle or a rectangle, for various values of  $\theta$ , including  $\theta = \pi/6, \pi/4, \pi/3, \pi/2$ .

(2) Prove that  $R_\theta$  is invertible and that its inverse is  $R_{-\theta}$ .

(3) For any two rotations  $R_\alpha$  and  $R_\beta$ , prove that

$$R_\beta \circ R_\alpha = R_\alpha \circ R_\beta = R_{\alpha+\beta}.$$

Use (2)-(3) to prove that the rotations in the plane form a commutative group denoted  $\text{SO}(2)$ .

**Problem B4 (110 pts).** Consider the affine map  $R_{\theta,(a_1,a_2)}$  in  $\mathbb{R}^2$  given by

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

(1) Prove that if  $\theta \neq k2\pi$ , with  $k \in \mathbb{Z}$ , then  $R_{\theta,(a_1,a_2)}$  has a unique fixed point  $(c_1, c_2)$ , that is, there is a unique point  $(c_1, c_2)$  such that

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = R_{\theta,(a_1,a_2)} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

and this fixed point is given by

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{1}{2 \sin(\theta/2)} \begin{pmatrix} \cos(\pi/2 - \theta/2) & -\sin(\pi/2 - \theta/2) \\ \sin(\pi/2 - \theta/2) & \cos(\pi/2 - \theta/2) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

(2) In this question, we still assume that  $\theta \neq k2\pi$ , with  $k \in \mathbb{Z}$ . By translating the coordinate system with origin  $(0, 0)$  to the new coordinate system with origin  $(c_1, c_2)$ , which means that if  $(x_1, x_2)$  are the coordinates with respect to the standard origin  $(0, 0)$  and if  $(x'_1, x'_2)$  are the coordinates with respect to the new origin  $(c_1, c_2)$ , we have

$$\begin{aligned} x_1 &= x'_1 + c_1 \\ x_2 &= x'_2 + c_2 \end{aligned}$$

and similarly for  $(y_1, y_2)$  and  $(y'_1, y'_2)$ , then show that

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = R_{\theta,(a_1,a_2)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

becomes

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

Conclude that with respect to the new origin  $(c_1, c_2)$ , the affine map  $R_{\theta, (a_1, a_2)}$  becomes the rotation  $R_\theta$ . We say that  $R_{\theta, (a_1, a_2)}$  is a *rotation of center*  $(c_1, c_2)$ .

(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  $k \in \mathbb{Z}$ ?

(4) Prove that the inverse of  $R_{\theta, (a_1, a_2)}$  is of the form  $R_{-\theta, (b_1, b_2)}$ , and find  $(b_1, b_2)$  in terms of  $\theta$  and  $(a_1, a_2)$ .

(5) Given two affine maps  $R_{\alpha, (a_1, a_2)}$  and  $R_{\beta, (b_1, b_2)}$ , prove that

$$R_{\beta, (b_1, b_2)} \circ R_{\alpha, (a_1, a_2)} = R_{\alpha + \beta, (t_1, t_2)}$$

for some  $(t_1, t_2)$ , and find  $(t_1, t_2)$  in terms of  $\beta$ ,  $(a_1, a_2)$  and  $(b_1, b_2)$ .

Even in the case where  $(a_1, a_2) = (0, 0)$ , prove that in general

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

Use (4)-(5) to show that the affine maps of the plane defined in this problem form a nonabelian group denoted  $\mathbf{SE}(2)$ .

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

**Problem B5 (80 pts).** A subset  $\mathcal{A}$  of  $\mathbb{R}^n$  is called an *affine subspace* if either  $\mathcal{A} = \emptyset$ , or there is some vector  $a \in \mathbb{R}^n$  and some subspace  $U$  of  $\mathbb{R}^n$  such that

$$\mathcal{A} = a + U = \{a + u \mid u \in U\}.$$

We define the dimension  $\dim(\mathcal{A})$  of  $\mathcal{A}$  as the dimension  $\dim(U)$  of  $U$ .

(1) If  $\mathcal{A} = a + U$ , why is  $a \in \mathcal{A}$ ?

What are affine subspaces of dimension 0? What are affine subspaces of dimension 1 (begin with  $\mathbb{R}^2$ )? What are affine subspaces of dimension 2 (begin with  $\mathbb{R}^3$ )?

Prove that any nonempty affine subspace is closed under affine combinations.

(2) Prove that if  $\mathcal{A} = a + U$  is any nonempty affine subspace, then  $\mathcal{A} = b + U$  for any  $b \in \mathcal{A}$ .

(3) Let  $\mathcal{A}$  be any nonempty subset of  $\mathbb{R}^n$  closed under affine combinations. For any  $a \in \mathcal{A}$ , prove that

$$U_a = \{x - a \in \mathbb{R}^n \mid x \in \mathcal{A}\}$$

is a (linear) subspace of  $\mathbb{R}^n$  such that

$$\mathcal{A} = a + U_a.$$

Prove that  $U_a$  does not depend on the choice of  $a \in \mathcal{A}$ ; that is,  $U_a = U_b$  for all  $a, b \in \mathcal{A}$ . In fact, prove that

$$U_a = U = \{y - x \in \mathbb{R}^n \mid x, y \in \mathcal{A}\}, \quad \text{for all } a \in \mathcal{A},$$

and so

$$\mathcal{A} = a + U, \quad \text{for any } a \in \mathcal{A}.$$

**Remark:** The subspace  $U$  is called the *direction* of  $\mathcal{A}$ .

(4) Two nonempty affine subspaces  $\mathcal{A}$  and  $\mathcal{B}$  are said to be *parallel* iff they have the same direction. Prove that if  $\mathcal{A} \neq \mathcal{B}$  and  $\mathcal{A}$  and  $\mathcal{B}$  are parallel, then  $\mathcal{A} \cap \mathcal{B} = \emptyset$ .

**Remark:** The above shows that affine subspaces behave quite differently from linear subspaces.

**Problem B6 (120 pts).** (Affine frames and affine maps) For any vector  $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ , let  $\widehat{v} \in \mathbb{R}^{n+1}$  be the vector  $\widehat{v} = (v_1, \dots, v_n, 1)$ . Equivalently,  $\widehat{v} = (\widehat{v}_1, \dots, \widehat{v}_{n+1}) \in \mathbb{R}^{n+1}$  is the vector defined by

$$\widehat{v}_i = \begin{cases} v_i & \text{if } 1 \leq i \leq n, \\ 1 & \text{if } i = n + 1. \end{cases}$$

(1) For any  $m + 1$  vectors  $(u_0, u_1, \dots, u_m)$  with  $u_i \in \mathbb{R}^n$  and  $m \leq n$ , prove that if the  $m$  vectors  $(u_1 - u_0, \dots, u_m - u_0)$  are linearly independent, then the  $m + 1$  vectors  $(\widehat{u}_0, \dots, \widehat{u}_m)$  are linearly independent.

(2) Prove that if the  $m + 1$  vectors  $(\widehat{u}_0, \dots, \widehat{u}_m)$  are linearly independent, then for 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.

Any  $m + 1$  vectors  $(u_0, u_1, \dots, u_m)$  such that the  $m + 1$  vectors  $(\widehat{u}_0, \dots, \widehat{u}_m)$  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  $\mathbb{R}^n$ .

(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  $(u_0, u_1, \dots, u_n)$ .

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  $(e_1, \dots, e_n)$ .

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  $(u_0, u_1, \dots, u_n)$ .

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  $v$ .

(4) Let  $(u_0, \dots, u_n)$  be any affine frame in  $\mathbb{R}^n$  and let  $(v_0, \dots, v_n)$  be any vectors in  $\mathbb{R}^m$ . Prove that there is a *unique* affine map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  such that

$$f(u_i) = v_i, \quad i = 0, \dots, n.$$

(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  $\mathbb{R}^n$ . Prove that 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

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

In the special case where  $(a_0, \dots, a_n)$  is the canonical affine frame with  $a_i = e_{i+1}$  for  $i = 0, \dots, n-1$  and  $a_n = (0, \dots, 0)$  (where  $e_i$  is the  $i$ th canonical basis vector), show that

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

and

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

For example, when  $n = 2$ , if we write  $b_i = (x_i, y_i)$ , then we have

$$A = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} x_1 - x_3 & x_2 - x_3 & x_3 \\ y_1 - y_3 & y_2 - y_3 & y_3 \\ 0 & 0 & 1 \end{pmatrix}.$$

(6) Recall that a nonempty affine subspace  $\mathcal{A}$  of  $\mathbb{R}^n$  is any nonempty subset of  $\mathbb{R}^n$  closed under affine combinations. For any affine map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , for any affine subspace  $\mathcal{A}$  of  $\mathbb{R}^n$ , and any affine subspace  $\mathcal{B}$  of  $\mathbb{R}^m$ , prove that  $f(\mathcal{A})$  is an affine subspace of  $\mathbb{R}^m$ , and that  $f^{-1}(\mathcal{B})$  is an affine subspace of  $\mathbb{R}^n$ .

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

(1) Prove that the  $k \times k$  matrix  $A^\top A$  and the matrix  $A$  have the same nullspace. Use this to prove that  $\text{rank}(A^\top A) = \text{rank}(A)$ . Similarly, prove that the  $n \times n$  matrix  $AA^\top$  and the matrix  $A^\top$  have the same nullspace, and conclude that  $\text{rank}(AA^\top) = \text{rank}(A^\top)$ .

We will prove later that  $\text{rank}(A^\top) = \text{rank}(A)$ .

(2) Let  $a_1, \dots, a_k$  be  $k$  linearly independent vectors in  $\mathbb{R}^n$  ( $1 \leq k \leq n$ ), and let  $A$  be the  $n \times k$  matrix whose  $i$ th column is  $a_i$ . Prove that  $A^\top A$  has rank  $k$ , and that it is invertible. Let  $P = A(A^\top A)^{-1}A^\top$  (an  $n \times n$  matrix). Prove that

$$\begin{aligned} P^2 &= P \\ P^\top &= P. \end{aligned}$$

What is the matrix  $P$  when  $k = 1$ ?

(3) Prove that the image of  $P$  is the subspace  $V$  spanned by  $a_1, \dots, a_k$ , or equivalently the set of all vectors in  $\mathbb{R}^n$  of the form  $Ax$ , with  $x \in \mathbb{R}^k$ . Prove that the nullspace  $U$  of  $P$  is the set of vectors  $u \in \mathbb{R}^n$  such that  $A^\top u = 0$ . Can you give a geometric interpretation of  $U$ ?

Conclude that  $P$  is a projection of  $\mathbb{R}^n$  onto the subspace  $V$  spanned by  $a_1, \dots, a_k$ , and that

$$\mathbb{R}^n = U \oplus V.$$

*Hint.* You may use results from HW2.

**TOTAL: 460 points.**