Fall 2020   CIS 515

Fundamentals of Linear Algebra and Optimization
Jean Gallier

Homework 3

October 05, 2020; Due October 19, 2020

Problem B1 (30 pts). A *rotation* R_θ 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_θ on a simple figure such as a triangle or a rectangle, for various values of θ, including $\theta = \pi/6, \pi/4, \pi/3, \pi/2$.

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

(3) For any two rotations R_α and R_β, 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 $\mathbf{SO}(2)$.

Problem B2 (100 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_θ. 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 θ 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 β, (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 B3 (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 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 B4 (120 pts). (Affine frames and affine maps) For any vector $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, let $\widehat{v} \in \mathbb{R}^{n+1}$ be the vector $\widehat{v} = (v_1, \ldots, v_n, 1)$. Equivalently, $\widehat{v} = (\widehat{v}_1, \ldots, \widehat{v}_{n+1}) \in \mathbb{R}^{n+1}$ 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 $(u_0, u_1, \ldots, u_m)$ with $u_i \in \mathbb{R}^n$ and $m \leq n$, prove that if the m vectors $(u_1 - u_0, \ldots, u_m - u_0)$ are linearly independent, then the $m+1$ vectors $(\widehat{u}_0, \ldots, \widehat{u}_m)$ are linearly independent.

(2) Prove that if the $m+1$ vectors $(\widehat{u}_0, \ldots, \widehat{u}_m)$ are linearly independent, then for any choice of i, with $0 \leq i \leq m$, the m vectors $u_j - u_i$ for $j \in \{0, \ldots, m\}$ 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) if $(u_0, u_1, \ldots, u_n)$ 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 $(\lambda_0, \lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n+1}$, with $\lambda_0 + \lambda_1 + \cdots + \lambda_n = 1$, such that

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

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

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

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

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

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

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

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

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

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

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

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

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

(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 $\mathbb{R}^n$. Prove that there is a unique $(n+1) \times (n+1)$ matrix

$$A = \begin{pmatrix} B & w \\ 0 & 1 \end{pmatrix}$$

corresponding to the unique affine map f such that

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

in the sense that

$$A\widehat{a}_i = \widehat{b}_i, \quad i = 0, \ldots, n,$$

and that A is given by

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

Make sure to prove that the bottom row of A is $(0, \ldots, 0, 1)$.

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

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

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

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

Hint. Write

$$\widehat{A} = \begin{pmatrix} \widehat{a_0} & \widehat{a_1} & \cdots & \widehat{a_n} \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & \ldots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

and

$$\widehat{B} = \begin{pmatrix} \widehat{b_0} & \widehat{b_1} & \cdots & \widehat{b_n} \end{pmatrix} = \begin{pmatrix} b_0 & b_1 & \ldots & b_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

We can write

$$A = \widehat{B}\widehat{A}^{-1} = \widehat{B}\mathcal{E}_n^{-1}\mathcal{E}_n\widehat{A}^{-1} = (\widehat{B}\mathcal{E}_n^{-1})(\widehat{A}\mathcal{E}_n^{-1})^{-1}.$$

The idea is to factor the unique affine map f that sends the affine frame $(a_0, \ldots, a_n)$ to $(b_0, \ldots, b_n)$ as the composition $f = f_2 \circ f_1$ of two unique affine maps f_1 and f_2, where f_1 maps the affine frame $(a_0, \ldots, a_n)$ to the canonical affine frame $(e_1, \ldots, e_n, e_0)$, and f_2 maps the the canonical affine frame $(e_1, \ldots, e_n, e_0)$ to $(b_0, \ldots, b_n)$. The inverse f_1^{-1} of f_1 is the unique affine map that sends the canonical affine frame $(e_1, \ldots, e_n, e_0)$ to the affine frame $(a_0, \ldots, a_n)$.

Prove that the set of $(n \times 1) \times (n + 1)$ matrices of the form

$$\begin{pmatrix} P & u \\ 0 & 1 \end{pmatrix},$$

where P is an invertible $n \times n$ matrix and $u \in \mathbb{R}^n$, is a group under matrix multiplication.

Another method goes as follows. Let $\mathcal{H}_{n+1}$ be the subset of $\mathbb{R}^{n+1}$ defined by

$$\mathcal{H}_{n+1} = \{\widehat{v} \mid v \in \mathbb{R}^n\} = \left\{ \begin{pmatrix} v \\ 1 \end{pmatrix} \mid v \in \mathbb{R}^n \right\}$$

called the *hyperplane of equation* $x_{n+1} = 1$. Check that $\mathcal{H}_{n+1}$ is an affine hyperplane with direction

$$\mathbb{R}^n \times \{0\} = \left\{ \begin{pmatrix} v \\ 0 \end{pmatrix} \mid v \in \mathbb{R}^n \right\}.$$

Prove that if an $(n + 1) \times (n + 1)$ matrix A of the form

$$A = \begin{pmatrix} B & w \\ 0 & 1 \end{pmatrix}$$

represents the unique affine map f from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

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

then

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

Prove that the last row of

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

is $(0, \ldots, 0, 1)$ (with n zeros). For this, prove the following two facts:

(1) The matrix A represents a linear map $\widehat{f}$ that maps the hyperplane $x_{n+1} = 1$ into the hyperplane $x_{n+1} = 1$,

(2) If A is a matrix representing a linear map $\widehat{f}$ from $\mathbb{R}^{n+1}$ to $\mathbb{R}^{n+1}$ and if $\widehat{f}$ maps the hyperplane $x_{n+1} = 1$ into the hyperplane $x_{n+1} = 1$, then the $(n+1)$th row of A is $(0, \ldots, 0, 1)$ (a row vector with n zeros).

(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 $f\colon \mathbb{R}^n \to \mathbb{R}^m$, 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 B5 (100 pts). (1) Let A be any invertible 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Prove that there is an invertible matrix S such that

$$SA = \begin{pmatrix} 1 & 0 \\ 0 & ad - bc \end{pmatrix},$$

where S is the product of at most four elementary matrices of the form $E_{i,j;\beta}$.

Conclude that every matrix A in $\mathbf{SL}(2)$ (the group of invertible 2×2 matrices A with $\det(A) = +1$) is the product of at most four elementary matrices of the form $E_{i,j;\beta}$.

For any $a \neq 0, 1$, give an explicit factorization as above for

$$A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}.$$

What is this decomposition for $a = -1$?

(2) Recall that a rotation matrix R (a member of the group $\mathbf{SO}(2)$) is a matrix of the form

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

Prove that if $\theta \neq k\pi$ (with $k \in \mathbb{Z}$), any rotation matrix can be written as a product

$$R = ULU,$$

where U is upper triangular and L is lower triangular of the form

$$U = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix}.$$

Therefore, every plane rotation (except a flip about the origin when $\theta = \pi$) can be written as the composition of three shear transformations!

(3) Recall that $E_{i,d}$ is the diagonal matrix

$$E_{i,d} = \mathrm{diag}(1, \ldots, 1, d, 1, \ldots, 1),$$

whose diagonal entries are all $+1$, except the (i, i)th entry which is equal to d.

Given any $n \times n$ matrix A, for any pair (i, j) of distinct row indices ($1 \leq i, j \leq n$), prove that there exist two elementary matrices $E_1(i, j)$ and $E_2(i, j)$ of the form $E_{k,\ell;\beta}$, such that

$$E_{j,-1} E_1(i, j) E_2(i, j) E_1(i, j) A = P(i, j) A,$$

the matrix obtained from the matrix A by permuting row i and row j. Equivalently, we have

$$E_1(i, j) E_2(i, j) E_1(i, j) A = E_{j,-1} P(i, j) A,$$

the matrix obtained from A by permuting row i and row j and multiplying row j by -1.

Prove that for every $i = 1, \ldots, n-1$, there exist four elementary matrices $E_3(i, d)$, $E_4(i, d)$, $E_5(i, d)$, $E_6(i, d)$ of the form $E_{k,\ell;\beta}$, such that

$$E_6(i, d) E_5(i, d) E_4(i, d) E_3(i, d) E_{n,d} = E_{i,d}.$$

What happens when $d = -1$, that is, what kind of simplifications occur?

Prove that all permutation matrices can be written as products of elementary operations of the form $E_{k,\ell;\beta}$ and the operation $E_{n,-1}$.

(4) Prove that for every invertible $n \times n$ matrix A, there is a matrix S such that

$$SA = \begin{pmatrix} I_{n-1} & 0 \\ 0 & d \end{pmatrix} = E_{n,d},$$

with $d = \det(A)$, and where S is a product of elementary matrices of the form $E_{k,\ell;\beta}$.

In particular, every matrix in $\mathbf{SL}(n)$ (the group of invertible $n \times n$ matrices A with $\det(A) = +1$) can be written as a product of elementary matrices of the form $E_{k,\ell;\beta}$. Prove that at most $n(n + 1) - 2$ such transformations are needed.

Extra Credit (20 points). Prove that every matrix in $\mathbf{SL}(n)$ can be written as a product of at most $(n-1)(\max\{n,3\}+1)$ elementary matrices of the form $E_{k,\ell;\beta}$.

Problem B6 (60 pts). Let E be a real vector space of dimension $n \geq 2$ and let F be any real vector space. Pick any basis $(u_1, \ldots, u_n)$ in E.

(1) Prove that for any bilinear alternating map $f\colon E \times E \to F$, for any two vectors $x = x_1u_1 + \cdots + x_nu_n$ and $y = y_1u_1 + \cdots + y_nu_n$, we have

$$f(x,y) = \sum_{1 \leq i < j \leq n} (x_iy_j - x_jy_i) f(u_i, u_j).$$

Observe that

$$x_iy_j - x_jy_i = \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}$$

is the determinant obtained from the $2 \times n$ matrix

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix}$$

by choosing two columns of index $i < j$ among the n columns.

Hint. Let $v = x_2u_2 + \cdots + x_nu_n$ and $w = y_2u_2 + \cdots + y_nu_n$. First prove that

$$\begin{aligned} f(x,y) = (x_1y_2 - x_2y_1)f(u_1,u_2) + (x_1y_3 - x_3y_1)f(u_1,u_3) + \cdots + (x_1y_n - x_ny_1)f(u_1,u_n) \\ + f(v,w). \end{aligned}$$

Then use induction.

(2) Prove that for any sequence $(w_{ij})_{1 \leq i < j \leq n}$ of $\binom{n}{2} = n(n-1)/2$ vectors $w_{ij} \in F$, there is a unique bilinear alternating map $f\colon E \times E \to F$ such that

$$f(u_i, u_j) = w_{ij}, \quad 1 \leq i < j \leq n,$$

and in fact,

$$f(x,y) = \sum_{1 \leq i < j \leq n} (x_iy_j - x_jy_i) w_{ij}.$$

Conclude that there is a bijection φ between the set $\mathrm{Alt}^2(E;F)$ of bilinear alternating maps $f\colon E \times E \to F$ and the product vector space $F^{n(n-1)/2}$ given by

$$\varphi(f) = (f(u_i, u_j))_{1 \leq i < j \leq n}.$$

Remark. Observe that when $F = \mathbb{R}$, if we let A be the $n \times n$ matrix given by $A = (f(e_i, e_j))$ and if we let X be the column vector with entries $(x_1, \ldots, x_n)$ and Y be the column vector with entries $(y_1, \ldots, y_n)$, then $A^\top = -A$ and $f(x,y) = X^\top AY$.

(3) We define addition and scalar multiplication on the set of bilinear alternating maps as follows. For any two bilinear alternating maps $f\colon E \times E \to F$ and $g\colon E \times E \to F$, for all $x, y \in E$ and all $\lambda \in \mathbb{R}$,

$$(f + g)(x, y) = f(x, y) + g(x, y),$$

and

$$(\lambda f)(x, y) = \lambda f(x, y).$$

Check (quickly) that $f + g$ and λf are bilinear and alternating, and that the set $\mathrm{Alt}^2(E; F)$ of bilinear alternating maps with the above addition and scalar multiplication is a real vector space.

(4) Prove that the bijection $\varphi\colon \mathrm{Alt}^2(E; F) \to F^{n(n-1)/2}$ in (2) given by

$$\varphi(f) = (f(u_i, u_j))_{1 \leq i < j \leq n}$$

is linear. Conclude that φ is an isomorphism of vector spaces, and that if F has dimension m, then $\mathrm{Alt}^2(E; F)$ has dimension $mn(n-1)/2$.

Extra Credit (50 pts).

(5) Let p be an integer such that $1 \leq p \leq n$. Consider the set $\mathrm{Alt}^p(E; F)$ of multilinear alternating maps $f\colon E^p \to F$. Prove that for any vectors $x_1, \ldots, x_p \in E$, if

$$x_i = x_{i1}u_1 + \cdots + x_{in}u_n, \quad i = 1, \ldots, p,$$

then

$$f(x_1, \ldots, x_p) = \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq n} \Delta_{j_1, j_2, \ldots, j_p}(x_1, \ldots, x_p) f(u_{j_1}, u_{j_2}, \ldots, u_{j_p}),$$

where $\Delta_{j_1, j_2, \ldots, j_p}(x_1, \ldots, x_p)$ is the determinant (of a $p \times p$ matrix)

$$\Delta_{j_1, j_2, \ldots, j_p}(x_1, \ldots, x_p) = \begin{vmatrix} x_{1j_1} & x_{1j_2} & \cdots & x_{1j_p} \\ x_{2j_1} & x_{2j_2} & \cdots & x_{2j_p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{pj_1} & x_{pj_2} & \cdots & x_{pj_p} \end{vmatrix}.$$

Observe that the above determinant is obtained from the $p \times n$ matrix

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix},$$

by choosing the columns of index $j_1, j_2, \ldots, j_p$ among the n columns.

Hint. First observe that

$$f(x_1, \ldots, x_p) = \sum_{(j_1, \ldots, j_p) \in \{1, \ldots, n\}^{\{1, \ldots, p\}}} x_{1j_1} \cdots x_{pj_p} f(u_{j_1}, \ldots, u_{j_p}),$$

where the sum extends over all sequences $(j_1, \ldots, j_p)$ of length p of elements from $\{1, \ldots, n\}$.

You will also need the fact that the notion of signature of a permutation, which was defined for permutations of the set $\{1, \ldots, n\}$, is defined in a similar way for permutations of the set $\{j_1, \ldots, j_p\}$, with $1 \leq j_1 < \cdots < j_p \leq n$.

(6) Give $\mathrm{Alt}^p(E; F)$ the structure of a vector space as in (3). Prove that the map $\varphi\colon \mathrm{Alt}^p(E; F) \to F^{\binom{n}{p}}$ given by

$$\varphi(f) = \big(f(u_{j_1}, u_{j_2}, \ldots, u_{j_p})\big)_{1 \leq j_1 < j_2 < \cdots < j_p \leq n}$$

is an isomorphism of vector spaces.

What more can you say when $p = n$? What is the dimension of $\mathrm{Alt}^n(E; F)$?

Suppose $F = \mathbb{R}$. Prove that the dimension of $\mathrm{Alt}^p(E; \mathbb{R})$ is $\binom{n}{p}$ (recall that $1 \leq p \leq n$). What is the dimension of $\mathrm{Alt}^n(E; \mathbb{R})$?

(7) Prove that for $p > n$, every multilinear alternating map $f\colon E^p \to F$ is the zero map.

TOTAL: 490 + 70 points.