Fall, 2013 CIS 515

Fundamentals of Linear Algebra and Optimization
Jean Gallier

Homework 6

November 26, 2013; Due December 10, 2013
Beginning of class

Problem B1 (20 $+\infty_2/\infty_1 \approx$ 40 pts). (1) Let H be the affine hyperplane in $\mathbb{R}^n$ given by the equation

$$a_1x_1 + \cdots + a_nx_n = c,$$

with $a_i \neq 0$ for some $i, 1 \leq i \leq n$. The linear hyperplane H_0 parallel to H is given by the equation

$$a_1x_1 + \cdots + a_nx_n = 0,$$

and we say that a vector $y \in \mathbb{R}^n$ is *orthogonal* (or *perpendicular*) to H iff y is orthogonal to H_0. Let h be the intersection of H with the line through the origin and perpendicular to H. Prove that the coordinates of h are given by

$$\frac{c}{a_1^2 + \cdots + a_n^2}(a_1, \ldots, a_n).$$

(2) For any point $p \in H$, prove that $\|h\| \leq \|p\|$. Thus, it is natural to define the *distance* $d(O, H)$ from the origin O to the hyperplane H as $d(O, H) = \|h\|$. Prove that

$$d(O, H) = \frac{|c|}{(a_1^2 + \cdots + a_n^2)^{\frac{1}{2}}}.$$

(3) Let S be a finite set of $n \geq 3$ points in the plane ($\mathbb{R}^2$). Prove that if for every pair of distinct points $p_i, p_j \in S$, there is a third point $p_k \in S$ (distinct from p_i and p_j) such that p_i, p_j, p_k belong to the same (affine) line, then all points in S belong to a common (affine) line.

Hint. Proceed by contradiction and use a minimality argument. This is either ∞-hard or relatively easy, depending how you proceed!

Problem B2 (10 pts). Let A be any real or complex $n \times n$ matrix and let $\| \ \|$ be any operator norm.

Prove that for every $m \geq 1$,

$$\|I\| + \sum_{k=1}^{m} \left\| \frac{A^k}{k!} \right\| \leq e^{\|A\|}.$$

If you know some analysis, deduce from the above that the sequence (E_m) of matrices

$$E_m = I + \sum_{k=1}^{m} \frac{A^k}{k!}$$

converges to a limit denoted e^A, and called the *exponential* of A.

Problem B3 (90 pts). (The space of closed polygons in $\mathbb{R}^2$, after Hausmann and Knutson)

An *open polygon* P in the plane is a sequence $P = (v_1, \ldots, v_{n+1})$ of point $v_i \in \mathbb{R}^2$ called *vertices* (with $n \geq 1$). A *closed polygon*, for short a *polygon*, is an open polygon $P = (v_1, \ldots, v_{n+1})$ such that $v_{n+1} = v_1$. The sequence of *edge vectors* $(e_1, \ldots, e_n)$ associated with the open (or closed) polygon $P = (v_1, \ldots, v_{n+1})$ is defined by

$$e_i = v_{i+1} - v_i, \quad i = 1, \ldots, n.$$

Thus, a closed or open polygon is also defined by a pair $(v_1, (e_1, \ldots, e_n))$, with the vertices given by

$$v_{i+1} = v_i + e_i, \quad i = 1, \ldots, n.$$

Observe that a polygon $(v_1, (e_1, \ldots, e_n))$ is closed iff

$$e_1 + \cdots + e_n = 0.$$

Since every polygon $(v_1, (e_1, \ldots, e_n))$ can be translated by $-v_1$, so that $v_1 = (0, 0)$, we may assume that our polygons are specified by a sequence of edge vectors.

Recall that the plane $\mathbb{R}^2$ is isomorphic to $\mathbb{C}$, via the isomorphism

$$(x, y) \mapsto x + iy.$$

We will represent each edge vector e_k by the square of a complex number $w_k = a_k + ib_k$. Thus, every sequence of complex numbers $(w_1, \ldots, w_n)$ defines a polygon (namely, $(w_1^2, \ldots, w_n^2)$). This representation is many-to-one: the sequences $(\pm w_1, \ldots, \pm w_n)$ describe the same polygon. To every sequence of complex numbers $(w_1, \ldots, w_n)$, we associate the pair of vectors (a, b), with $a, b \in \mathbb{R}^n$, such that if $w_k = a_k + ib_k$, then

$$a = (a_1, \ldots, a_n), \quad b = (b_1, \ldots, b_n).$$

The mapping

$$(w_1, \ldots, w_n) \mapsto (a, b)$$

is clearly a bijection, so we can also represent polygons by pairs of vectors $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$.

(a) Prove that a polygon P represented by a pair of vectors $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$ is closed iff $a \cdot b = 0$ and $\|a\|_2 = \|b\|_2$.

(b) Given a polygon P represented by a pair of vectors $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$, the length $l(P)$ of the polygon P is defined by $l(P) = |w_1|^2 + \cdots + |w_n|^2$, with $w_k = a_k + ib_k$. Prove that

$$l(P) = \|a\|_2^2 + \|b\|_2^2 .$$

Deduce from (a) and (b) that every closed polygon of length 2 with n edges is represented by a $n \times 2$ matrix A such that $A^\top A = I$.

Remark: The space of all a $n \times 2$ real matrices A such that $A^\top A = I$ is a space known as the *Stiefel manifold* $S(2, n)$.

(c) Recall that in $\mathbb{R}^2$, the rotation of angle θ specified by the matrix

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

is expressed in terms of complex numbers by the map

$$z \mapsto ze^{i\theta}.$$

Let P be a polygon represented by a pair of vectors $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$. Prove that the polygon $R_\theta(P)$ obtained by applying the rotation R_θ to every vertex $w_k^2 = (a_k + ib_k)^2$ of P is specified by the pair of vectors

$$(\cos(\theta/2)a - \sin(\theta/2)b,\ \sin(\theta/2)a + \cos(\theta/2)b) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}.$$

(d) The reflection ρ_x about the x-axis corresponds to the map

$$z \mapsto \overline{z},$$

whose matrix is,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Prove that the polygon $\rho_x(P)$ obtained by applying the reflection ρ_x to every vertex $w_k^2 = (a_k + ib_k)^2$ of P is specified by the pair of vectors

$$(a, -b) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(e) Let $Q \in \mathbf{O}(2)$ be any isometry such that $\det(Q) = -1$ (a reflection). Prove that there is a rotation $R_{-\theta} \in \mathbf{SO}(2)$ such that

$$Q = \rho_x \circ R_{-\theta}.$$

Prove that the isometry Q, which is given by the matrix

$$Q = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},$$

is the reflection about the line corresponding to the angle $\theta/2$ (the line of equation $y = \tan(\theta/2)x$).

Prove that the polygon $Q(P)$ obtained by applying the reflection $Q = \rho_x \circ R_{-\theta}$ to every vertex $w_k^2 = (a_k + ib_k)^2$ of P, is specified by the pair of vectors

$$(\cos(\theta/2)a + \sin(\theta/2)b,\ \sin(\theta/2)a - \cos(\theta/2)b) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ \sin(\theta/2) & -\cos(\theta/2) \end{pmatrix}.$$

(f) Define an equivalence relation $\sim$ on $S(2, n)$ such that if $A_1, A_2 \in S(2, n)$ are any $n \times 2$ matrices such that $A_1^\top A_1 = A_2^\top A_2 = I$, then

$$A_1 \sim A_2 \quad \text{iff} \quad A_2 = A_1 Q \quad \text{for some } Q \in \mathbf{O}(2).$$

Prove that the quotient $G(2, n) = S(2, n)/\sim$ is in bijection with the set of all 2-dimensional subspaces (the planes) of $\mathbb{R}^n$. The space $G(2, n)$ is called a *Grassmannian manifold*.

Prove that up to translations and isometries in $\mathbf{O}(2)$ (rotations and reflections), the n-sided closed polygons of length 2 are represented by planes in $G(2, n)$.

Problem B4 (100 pts). (a) For any matrix

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

if we let $\theta = \sqrt{a^2 + b^2 + c^2}$ and

$$B = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix},$$

prove that

$$\begin{aligned} A^2 &= -\theta^2 I + B, \\ AB &= BA = 0. \end{aligned}$$

From the above, deduce that

$$A^3 = -\theta^2 A.$$

(b) Prove that the exponential map $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is given by

$$\exp A = e^A = \cos\theta\, I_3 + \frac{\sin\theta}{\theta} A + \frac{(1-\cos\theta)}{\theta^2} B,$$

or, equivalently, by

$$e^A = I_3 + \frac{\sin\theta}{\theta} A + \frac{(1-\cos\theta)}{\theta^2} A^2, \quad \text{if } \theta \neq 0,$$

with $\exp(0_3) = I_3$.

(c) Prove that e^A is an orthogonal matrix of determinant $+1$, i.e., a rotation matrix.

(d) Prove that the exponential map $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is surjective. For this, proceed as follows: Pick any rotation matrix $R \in \mathbf{SO}(3)$;

(1) The case $R = I$ is trivial.

(2) If $R \neq I$ and $\mathrm{tr}(R) \neq -1$, then

$$\exp^{-1}(R) = \left\{ \frac{\theta}{2\sin\theta}(R - R^T) \;\middle|\; 1 + 2\cos\theta = \mathrm{tr}(R) \right\}.$$

(Recall that $\mathrm{tr}(R) = r_{1\,1} + r_{2\,2} + r_{3\,3}$, the *trace* of the matrix R).

Show that there is a unique skew-symmetric B with corresponding θ satisfying $0 < \theta < \pi$ such that $e^B = R$.

(3) If $R \neq I$ and $\mathrm{tr}(R) = -1$, then prove that the eigenvalues of R are $1, -1, -1$, that $R = R^\top$, and that $R^2 = I$. Prove that the matrix

$$S = \frac{1}{2}(R - I)$$

is a symmetric matrix whose eigenvalues are $-1, -1, 0$. Thus, S can be diagonalized with respect to an orthogonal matrix Q as

$$S = Q \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^\top.$$

Prove that there exists a skew symmetric matrix

$$U = \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}$$

so that

$$U^2 = S = \frac{1}{2}(R - I).$$

Observe that

$$U^2 = \begin{pmatrix} -(c^2+d^2) & bc & bd \\ bc & -(b^2+d^2) & cd \\ bd & cd & -(b^2+c^2) \end{pmatrix},$$

and use this to conclude that if $U^2 = S$, then $b^2 + c^2 + d^2 = 1$. Then, show that

$$\exp^{-1}(R) = \left\{ (2k+1)\pi \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}, \ k \in \mathbb{Z} \right\},$$

where (b, c, d) is any unit vector such that for the corresponding skew symmetric matrix U, we have $U^2 = S$.

(e) To find a skew symmetric matrix U so that $U^2 = S = \frac{1}{2}(R - I)$ as in (d), we can solve the system

$$\begin{pmatrix} b^2-1 & bc & bd \\ bc & c^2-1 & cd \\ bd & cd & d^2-1 \end{pmatrix} = S.$$

We immediately get b^2, c^2, d^2, and then, since one of b, c, d is nonzero, say b, if we choose the positive square root of b^2, we can determine c and d from bc and bd.

Implement a computer program to solve the above system.

Problem B5 (120 pts). (a) Consider the set of affine maps ρ of $\mathbb{R}^3$ defined such that

$$\rho(X) = \alpha R X + W,$$

where R is a rotation matrix (an orthogonal matrix of determinant $+1$), W is some vector in $\mathbb{R}^3$, and $\alpha \in \mathbb{R}$ with $\alpha > 0$. Every such a map can be represented by the 4×4 matrix

$$\begin{pmatrix} \alpha R & W \\ 0 & 1 \end{pmatrix}$$

in the sense that

$$\begin{pmatrix} \rho(X) \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha R & W \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix}$$

iff

$$\rho(X) = \alpha R X + W.$$

Prove that these maps form a group, denoted by $\mathbf{SIM}(3)$ (the *direct affine similitudes* of $\mathbb{R}^3$).

(b) Let us now consider the set of 4×4 real matrices of the form

$$B = \begin{pmatrix} \Gamma & W \\ 0 & 0 \end{pmatrix},$$

where Γ is a matrix of the form

$$\Gamma = \lambda I_3 + \Omega,$$

with

$$\Omega = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

so that

$$\Gamma = \begin{pmatrix} \lambda & -c & b \\ c & \lambda & -a \\ -b & a & \lambda \end{pmatrix},$$

and W is a vector in $\mathbb{R}^3$.

Verify that this set of matrices is a vector space isomorphic to $(\mathbb{R}^7, +)$. This vector space is denoted by $\mathfrak{sim}(3)$.

(c) Given a matrix

$$B = \begin{pmatrix} \Gamma & W \\ 0 & 0 \end{pmatrix}$$

as in (b), prove that

$$B^n = \begin{pmatrix} \Gamma^n & \Gamma^{n-1}W \\ 0 & 0 \end{pmatrix}$$

where $\Gamma^0 = I_3$. Prove that

$$e^B = \begin{pmatrix} e^\Gamma & VW \\ 0 & 1 \end{pmatrix},$$

where

$$V = I_3 + \sum_{k \geq 1} \frac{\Gamma^k}{(k+1)!}.$$

(d) Prove that if $\Gamma = \lambda I_3 + \Omega$ as in (b), then

$$V = I_3 + \sum_{k \geq 1} \frac{\Gamma^k}{(k+1)!} = \int_0^1 e^{\Gamma t} dt.$$

(e) For any matrix $\Gamma = \lambda I_3 + \Omega$, with

$$\Omega = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

if we let $\theta = \sqrt{a^2 + b^2 + c^2}$, then prove that

$$e^{\Gamma} = e^{\lambda} e^{\Omega} = e^{\lambda}\left(I_3 + \frac{\sin\theta}{\theta}\Omega + \frac{(1-\cos\theta)}{\theta^2}\Omega^2 \right), \quad \text{if } \theta \neq 0,$$

and $e^{\Gamma} = e^{\lambda} I_3$ if $\theta = 0$.

Hint. You may use the fact that if $AB = BA$, then $e^{A+B} = e^A e^B$. In general, $e^{A+B} \neq e^A e^B$!

(f) Prove that

1. If $\theta = 0$ and $\lambda = 0$, then
$$V = I_3.$$
2. If $\theta = 0$ and $\lambda \neq 0$, then
$$V = \frac{(e^{\lambda} - 1)}{\lambda} I_3;$$
3. If $\theta \neq 0$ and $\lambda = 0$, then
$$V = I_3 + \frac{(1-\cos\theta)}{\theta^2}\Omega + \frac{(\theta - \sin\theta)}{\theta^3}\Omega^2.$$
4. If $\theta \neq 0$ and $\lambda \neq 0$, then
$$\begin{aligned} V = {} & \frac{(e^{\lambda}-1)}{\lambda} I_3 + \frac{(\theta(1 - e^{\lambda}\cos\theta) + e^{\lambda}\lambda\sin\theta)}{\theta(\lambda^2+\theta^2)}\Omega \\ & + \left(\frac{(e^{\lambda}-1)}{\lambda\theta^2} - \frac{e^{\lambda}\sin\theta}{\theta(\lambda^2+\theta^2)} - \frac{\lambda(e^{\lambda}\cos\theta - 1)}{\theta^2(\lambda^2+\theta^2)} \right)\Omega^2. \end{aligned}$$

Hint. You will need to compute $\int_0^1 e^{\lambda t}\sin\theta t\, dt$ and $\int_0^1 e^{\lambda t}\cos\theta t\, dt$.

(g) Prove that V is invertible iff $\lambda \neq 0$ or $\theta \neq k2\pi$, with $k \in \mathbb{Z} - \{0\}$.

Hint. Express the eigenvalues of V in terms of the eigenvalues of Γ.

In the special case where $\lambda = 0$, show that

$$V^{-1} = I - \frac{1}{2}\Omega + \frac{1}{\theta^2}\left(1 - \frac{\theta\sin\theta}{2(1-\cos\theta)}\right)\Omega^2, \quad \text{if } \theta \neq 0.$$

Hint. Assume that the inverse of V is of the form

$$Z = I_3 + a\Omega + b\Omega^2,$$

and show that a, b, are given by a system of linear equations that always has a unique solution.

(h) Prove that the exponential map $\exp\colon \mathfrak{sim}(3) \to \mathbf{SIM}(3)$, given by $\exp(B) = e^B$, is surjective. You may use the fact that $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is surjective, proved in another Problem.

Remark: As in the case of the plane, curves in $\mathbf{SIM}(3)$ can be used to describe certain deformations of bodies in $\mathbb{R}^3$.

Problem B6 (40 pts). (1) Consider the matrix

$$A = \begin{pmatrix} 0 & 0 & -a_3 \\ 1 & 0 & -a_2 \\ 0 & 1 & -a_1 \end{pmatrix}.$$

Prove that the characteristic polynomial $\chi_A(z) = \det(zI - A)$ of A is given by

$$\chi_A(z) = z^3 + a_1 z^2 + a_2 z + a_3.$$

(2) Consider the matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & -a_4 \\ 1 & 0 & 0 & -a_3 \\ 0 & 1 & 0 & -a_2 \\ 0 & 0 & 1 & -a_1 \end{pmatrix}.$$

Prove that the characteristic polynomial $\chi_A(z) = \det(zI - A)$ of A is given by

$$\chi_A(z) = z^4 + a_1 z^3 + a_2 z^2 + a_3 z + a_4.$$

(3) Consider the $n \times n$ matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & -a_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & -a_2 \\ 0 & 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}.$$

Prove that the characteristic polynomial $\chi_A(z) = \det(zI - A)$ of A is given by

$$\chi_A(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_{n-1} z + a_n.$$

Hint. Use induction.

Explain why finding the roots of a polynomial (with real or complex coefficients), and finding the eigenvalues of a (real or complex) matrix, are equivalent problems, in the sense that if we have a method for solving one of these problems, then we have a method to solve the other.

TOTAL: 420 points.