Spring, 2005 CIS 610

Advanced Geometric Methods in Computer Science

Jean Gallier

Homework 1

February 2, 2005; Due February 21, 2005

"A problems" are for practice only, and should not be turned in.

Problem A1. Given a finite dimensional Euclidean space, E, if U and V are two orthogonal subspaces that span E, i.e., $E = U \oplus V$, we have the linear projections $p_U \colon E \to U$ and $p_V \colon E \to V$. Recall: since every $w \in E$ can be written uniquely as $w = u + v$, with $u \in U$ and $v \in V$, we have $p_U(w) = u$, $p_V(w) = v$ and $p_U(w) + p_V(w) = w$, for all $w \in E$. We define the *orthogonal reflection with respect to U and parallel to V* as the linear map, s, given by

$$s(w) = 2p_U(w) - w = w - 2p_V(w),$$

for all $w \in E$. Observe that $s \circ s = \mathrm{id}$, that s is the identity on U and $s = -\mathrm{id}$ on V. When $U = H$ is a hyperplane, s is called a *hyperplane reflection* (about H).

(a) If w is any nonzero vector orthogonal to the hyperplane H, prove that s is given by

$$s(x) = x - 2\frac{\langle x, w\rangle}{\|w\|^2} w,$$

for all $x \in E$. (Here, $\|w\|^2 = \langle w, w\rangle$.)

(b) In matrix form, if the vector w is represented by the column vector W, show that the matrix of the hyperplane reflection about the hyperplane $K = \{w\}^\perp$ is

$$I - 2\frac{WW^\top}{W^\top W}.$$

Such matrices are called *Householder matrices*.

Problem A2. As in A1, let E be a finite dimensional Euclidean space and assume E is nontrivial, i.e, $\dim(E) \geq 1$. Prove that if $u, v \in E$ are any two nonzero vectors and $\|u\| = \|v\|$, then there is a hyperplane, H, so that the reflection s about H sends u to v ($v = s(u)$) and if $u \neq v$, then this reflection is unique.

Problem A3. Given a finite dimensional Euclidean space, E, recall that a linear map, $f \colon E \to E$, is an *isometry* iff

$$\langle f(u), f(v)\rangle = \langle u, v\rangle, \quad \text{for all } u, v \in E.$$

(a) Prove that a linear map, f, is an isometry iff

$$f^* \circ f = f \circ f^* = \mathrm{id},$$

where f^* denotes the adjoint of f.

(b) Note that an isometry, f, also preserves the Euclidean norm $\|u\| = \sqrt{\langle u, u\rangle}$, i.e., $\|f(u)\| = \|u\|$, for all $u \in E$.

Is the following converse true: If f is a linear map and $\|f(u)\| = \|u\|$, for all $u \in E$, then f is an isometry?

Is this statement still true if we do not assume that f is linear?

(c) For *any* map, $f\colon E \to E$, show that the condition

$$\langle f(u), f(v)\rangle = \langle u, v\rangle, \quad \text{for all } u, v \in E$$

implies that f is actually linear (remember, E has finite dimension).

Problem A4. Let E and F be normed vector spaces (as defined in the transparencies, Section 3.1).

(a) Check that

$$\sup_{u \neq 0} \frac{\|Au\|}{\|u\|} = \sup_{\|u\|=1} \|Au\|.$$

Hint. Use property (b) of a norm.

(b) Prove that a linear map, $A\colon E \to F$, is bounded iff it is linear. (Again, property (b) of norms will be useful.)

(c) Prove that every norm on $\mathbb{R}^n$ or $\mathbb{C}^n$ is continuous.

(d) Two norms $\| \|_1$ and $\| \|_2$ are *equivalent* iff there exist $c_1, c_2 > 0$ so that $\|u\|_1 \leq c_1 \|u\|_2$ and $\|u\|_2 \leq c_2 \|u\|_1$, for all $u \in E$. Prove that on a finite dimensional vector space, E, any two norms are equivalent.

Hint. If E is finite-dimensional, then E is isomorphic to $\mathbb{R}^n$ or to $\mathbb{C}^n$. Use the fact that the $(n-1)$-sphere

$$S^{n-1} = \{u \in E \mid \|u\|_2 = 1\}$$

is compact and consider the values of the functions $x \mapsto \frac{\|x\|_1}{\|x\|_2}$ and $x \mapsto \frac{\|x\|_2}{\|x\|_1}$ on S^{n-1}.

(e) Use (d) to prove that if E is finite-dimensional, then every linear map, $A\colon E \to F$, is bounded (E and F are normed vector spaces).

Problem A5. Prove Proposition 3.1.6 of the transparencies.

Problem A6. Given an $m \times n$ matrix, A, prove that its Frobenius norm,

$$\|A\|_F = \sqrt{\sum_{ij} |a_{i\,j}|^2}$$

satisfies

$$\|A\|_F = \sqrt{\mathrm{tr}(A^* A)} = \sqrt{\mathrm{tr}(AA^*)}$$

where $\mathrm{tr}(B)$ is the trace of the square matrix B (the sum of its diagonal elements).

"B problems" must be turned in.

Problem B1 (30 pts). Prove Proposition 3.1.7:

Let A be an $m \times n$ matrix (over $\mathbb{R}$ or $\mathbb{C}$) and let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_p$ be its singular values (where $p = \min(m, n)$). Then, the following properties hold:

1. $\|Au\| \leq \|A\| \, \|u\|$, *where $\|A\|$ is a subordinate norm and* $\|Au\|_2 \leq \|A\|_F \|u\|_2$, *where $\|A\|_F$ is the Frobenius norm.*
2. $\|AB\| \leq \|A\| \, \|B\|$, *for a subordinate norm or the Frobenius norm.*
3. $\|UAV\| = \|A\|$, *if U and V are orthogonal (or unitary) and $\| \ \|$ is the Frobenius norm or the subordinate norm $\| \ \|_2$.*
4. $\|A\|_\infty = \max_i \sum_j |a_{i\,j}|$.
5. $\|A\|_1 = \max_j \sum_i |a_{i\,j}|$.
6. $\|A\|_2 = \mu_1 = \sqrt{\lambda_{\max}(A^* A)}$, *where $\lambda_{\max}(A^* A)$ is the largest eigenvalue of $A^* A$.*
7. $\|A\|_F = \sqrt{\sum_{i=1}^{p} \mu_i^2}$, *where $p = \min(m, n)$.*
8. $\|A\|_2 \leq \|A\|_F \leq \sqrt{p}\, \|A\|_2$.

In (4), (5), (6), (8), the matrix norms are the subordinate norms induced by the corresponding norms ($\| \ \|_\infty$, $\| \ \|_1$ and $\| \ \|_2$) on $\mathbb{R}^m$ and $\mathbb{R}^n$.

Problem B2 (30 pts). Prove Proposition 3.1.8:

Let A be an $m \times n$ matrix of rank r and let $VDU^\top = A$ be an SVD for A. Write u_i for the columns of U, v_i for the columns of V and $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_p$ for the singular values of A ($p = \min(m, n)$). Then, a matrix of rank $k < r$ closest to A (in the $\| \ \|_2$ norm) is given by

$$A_k = \sum_{i=1}^{k} \mu_i v_i u_i^\top = V \mathrm{diag}(\mu_1, \ldots, \mu_k) U^\top$$

and $\|A - A_k\|_2 = \mu_{k+1}$.

Hint. You will need to prove that if B is any rank k matrix ($k < r$) then $\|A - B\|_2 \geq \mu_{k+1}$. Since $\mathrm{rk}(B) = k$, the kernel of B has dimension $n - k$. Note that the space spanned by $u_1, \ldots, u_{k+1}$ has dimension $k + 1$; deduce that there must be a unit vector, h, in their intersection and use

$$\|A - B\|_2 \geq \|(A - B)h\|_2 .$$

Problem B3 (40 pts). (a) Prove Lemma 3.2.2:

If B is a symmetric positive semi-definite $d \times d$ matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0$ and associated eigenvectors $u_1, \ldots, u_d$, then

$$\max_{x \neq 0} \frac{x^\top B x}{x^\top x} = \lambda_1$$

(with the maximum attained for $x = u_1$) and

$$\max_{x \neq 0, x \in \{u_1, \ldots, u_k\}^\perp} \frac{x^\top B x}{x^\top x} = \lambda_{k+1}$$

(with the maximum attained for $x = u_{k+1}$), where $1 \leq k \leq d - 1$.

(b) Prove Theorem 3.2.3:

Let X be an $n \times d$ matrix of data points, $X_1, \ldots, X_n$, and let μ be the centroid of the X_i's. If $X - \mu = VDU^\top$ is an SVD decomposition of $X - \mu$ and if the main diagonal of D consists of the singular values $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_d$, then a kth principal component of X is given by

$$Y_k = (X - \mu)u_k = k\text{th column of } VD,$$

where u_k is the kth column of U. Furthermore,

$$\mathrm{var}(Y_k) = \frac{\mu_k^2}{n-1}$$

and $\mathrm{cov}(Y_h, Y_k) = 0$, whenever $h \neq k$.

Problem B4 (50 pts). The purpose of this problem is to prove that given any self-adjoint linear map $f \colon E \to E$ (i.e., such that $f^* = f$), where E is a Euclidean space of dimension $n \geq 3$, given an orthonormal basis $(e_1, \ldots, e_n)$, there are $n - 2$ isometries h_i, hyperplane reflections or the identity, such that the matrix of

$$h_{n-2} \circ \cdots \circ h_1 \circ f \circ h_1 \circ \cdots \circ h_{n-2}$$

is a symmetric tridiagonal matrix.

(1) Prove that for any isometry $f \colon E \to E$, we have $f = f^* = f^{-1}$ iff $f \circ f = \mathrm{id}$.

Prove that if f and h are self-adjoint linear maps ($f^* = f$ and $h^* = h$), then $h \circ f \circ h$ is a self-adjoint linear map.

(2) Proceed by induction, taking inspiration from the proof of the triangular decomposition given in Lemma 7.3.1 of my book, *Geometric Methods and Applications*. Let V_k be the subspace spanned by $(e_{k+1}, \ldots, e_n)$. For the base case, proceed as follows.

Let

$$f(e_1) = a_1^0 e_1 + \cdots + a_n^0 e_n,$$

and let

$$r_{1,2} = \left\| a_2^0 e_2 + \cdots + a_n^0 e_n \right\|.$$

Find an isometry h_1 (reflection or id) such that

$$h_1(f(e_1) - a_1^0 e_1) = r_{1,2}\, e_2.$$

Observe that

$$w_1 = r_{1,2}\, e_2 + a_1^0 e_1 - f(e_1) \in V_1,$$

and prove that $h_1(e_1) = e_1$, so that

$$h_1 \circ f \circ h_1(e_1) = a_1^0 e_1 + r_{1,2}\, e_2.$$

Let $f_1 = h_1 \circ f \circ h_1$.

Assuming by induction that

$$f_k = h_k \circ \cdots \circ h_1 \circ f \circ h_1 \circ \cdots \circ h_k$$

has a tridiagonal matrix up to the kth row and column, $1 \leq k \leq n-3$, let

$$f_k(e_{k+1}) = a_k^k e_k + a_{k+1}^k e_{k+1} + \cdots + a_n^k e_n,$$

and let

$$r_{k+1,\,k+2} = \left\| a_{k+2}^k e_{k+2} + \cdots + a_n^k e_n \right\|.$$

Find an isometry h_{k+1} (reflection or id) such that

$$h_{k+1}(f_k(e_{k+1}) - a_k^k e_k - a_{k+1}^k e_{k+1}) = r_{k+1,\,k+2}\, e_{k+2}.$$

Observe that

$$w_{k+1} = r_{k+1,\,k+2}\, e_{k+2} + a_k^k e_k + a_{k+1}^k e_{k+1} - f_k(e_{k+1}) \in V_{k+1},$$

and prove that $h_{k+1}(e_k) = e_k$ and $h_{k+1}(e_{k+1}) = e_{k+1}$, so that

$$h_{k+1} \circ f_k \circ h_{k+1}(e_{k+1}) = a_k^k e_k + a_{k+1}^k e_{k+1} + r_{k+1,\,k+2}\, e_{k+2}.$$

Let $f_{k+1} = h_{k+1} \circ f_k \circ h_{k+1}$, and finish the proof.

Do f and f_{n-2} have the same eigenvalues? If so, explain why.

(3) Prove that given any symmetric $n \times n$-matrix A, there are $n-2$ matrices $H_1, \ldots, H_{n-2}$, Householder matrices or the identity, such that

$$B = H_{n-2} \cdots H_1 A H_1 \cdots H_{n-2}$$

is a symmetric tridiagonal matrix.

Problem B5 (20 pts). As discussed in class, let X be the 10×2 centered matrix consisting of the year of birth and the length of beard of our ten mathematicians:

Name	year	length
Carl Friedrich Gauss	-51.4	-5.6
Camille Jordan	9.6	6.4
Adrien-Marie Legendre	-76.4	-5.6
Bernhard Riemann	-2.4	9.4
David Hilbert	33.6	-3.6
Henri Poincaré	25.6	-0.6
Emmy Noether	53.6	-5.6
Karl Weierstrass	13.4	-5.6
Eugenio Beltrami	6.6	-3.6
Hermann Schwarz	14.6	14.4

Compute the principal directions and the two PC's of X. Plot your results (You may use *Matlab*, *Mathematica*, etc.). Can you conclude anything?

Problem B6 (40 pts). (a) Given a rotation matrix

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

where $0 < \theta < \pi$, prove that there is a skew symmetric matrix B such that

$$R = (I - B)(I + B)^{-1}.$$

(b) If B is a skew symmetric $n \times n$ matrix, prove that $\lambda I_n - B$ and $\lambda I_n + B$ are invertible for all $\lambda \neq 0$, and that they commute.

(c) Prove that

$$R = (\lambda I_n - B)(\lambda I_n + B)^{-1}$$

is a rotation matrix that does not admit -1 as an eigenvalue. (Recall, a rotation is an orthogonal matrix R with positive determinant, i.e., $\det(R) = 1$.)

(d) Given any rotation matrix R that does not admit -1 as an eigenvalue, prove that there is a skew symmetric matrix B such that

$$R = (I_n - B)(I_n + B)^{-1} = (I_n + B)^{-1}(I_n - B).$$

This is known as the *Cayley representation* of rotations (Cayley, 1846).

(e) Given any rotation matrix R, prove that there is a skew symmetric matrix B such that

$$R = \left((I_n - B)(I_n + B)^{-1}\right)^2.$$

Problem B7 (20 pts). Given any hyperplane, H, in $\mathbb{R}^m$ and given any point, $x \in \mathbb{R}^m$, the distance from x to H is defined by

$$d(x, H) = \min_{h \in H} d(x, h),$$

where $d(x, h)$ is the usual Euclidean distance in $\mathbb{R}^m$.

(a) If the hyperplane, H, is given by the equation

$$a_1 x_1 + \cdots + a_m x_m + c = 0,$$

then prove that

$$d(x, H) = \frac{|a_1 x_1 + \cdots + a_m x_m + c|}{\|a\|},$$

where $a = (a_1, \ldots, a_m) \neq 0$ and $x = (x_1, \ldots, x_m)$.

(b) Given a data set of n points, $X_1, \ldots, X_n \in \mathbb{R}^d$, prove that a hyperplane, H, that best approximates $X_1, \ldots, X_n$ in the least squares sense is a hyperplane that minimizes the sum of the square distances of each X_i to H.

Extra Credit (40 pts). Write a computer program implementing the method of Problem B4(3).

TOTAL: 230 + 40 points.