Fall, 2014   CIS 515

Fundamentals of Linear Algebra and Optimization

Jean Gallier

Homework 5

November 6, 2014; Due November 25 2014
Beginning of classes

Problem B1 (60 pts). (1) Let A be any $n \times n$ matrix such that the sum of the entries of every row of A is the same (say c_1), and the sum of entries of every column of A is the same (say c_2). Prove that $c_1 = c_2$.

(2) Prove that for any $n \geq 2$, the $2n - 2$ equations asserting that the sum of the entries of every row of A is the same, and the sum of entries of every column of A is the same are lineary independent. For example, when $n = 4$, we have the following 6 equations

$$\begin{aligned}
a_{11} + a_{12} + a_{13} + a_{14} - a_{21} - a_{22} - a_{23} - a_{24} &= 0 \\
a_{21} + a_{22} + a_{23} + a_{24} - a_{31} - a_{32} - a_{33} - a_{34} &= 0 \\
a_{31} + a_{32} + a_{33} + a_{34} - a_{41} - a_{42} - a_{43} - a_{44} &= 0 \\
a_{11} + a_{21} + a_{31} + a_{41} - a_{12} - a_{22} - a_{32} - a_{42} &= 0 \\
a_{12} + a_{22} + a_{32} + a_{42} - a_{13} - a_{23} - a_{33} - a_{43} &= 0 \\
a_{13} + a_{23} + a_{33} + a_{43} - a_{14} - a_{24} - a_{34} - a_{44} &= 0.
\end{aligned}$$

Hint. Group the equations as above; that is, first list the $n - 1$ equations relating the rows, and then list the $n - 1$ equations relating the columns. Prove that the first $n - 1$ equations are linearly independent, and that the last $n - 1$ equations are also linearly independent. Then, find a relationship between the two groups of equations that will allow you to prove that they span subspace V^r and V^c such that $V^r \cap V^c = (0)$.

(3) Now consider *magic squares*. Such matrices satisfy the two conditions about the sum of the entries in each row and in each column to be the same number, and also the additional two constraints that the main descending and the main ascending diagonals add up to this common number. Traditionally, it is also required that the entries in a magic square are positive integers, but we will consider generalized magic square with arbitrary real entries. For example, in the case $n = 4$, we have the following system of 8 equations:

$$
\begin{aligned}
a_{11} + a_{12} + a_{13} + a_{14} - a_{21} - a_{22} - a_{23} - a_{24} &= 0 \\
a_{21} + a_{22} + a_{23} + a_{24} - a_{31} - a_{32} - a_{33} - a_{34} &= 0 \\
a_{31} + a_{32} + a_{33} + a_{34} - a_{41} - a_{42} - a_{43} - a_{44} &= 0 \\
a_{11} + a_{21} + a_{31} + a_{41} - a_{12} - a_{22} - a_{32} - a_{42} &= 0 \\
a_{12} + a_{22} + a_{32} + a_{42} - a_{13} - a_{23} - a_{33} - a_{43} &= 0 \\
a_{13} + a_{23} + a_{33} + a_{43} - a_{14} - a_{24} - a_{34} - a_{44} &= 0 \\
a_{22} + a_{33} + a_{44} - a_{12} - a_{13} - a_{14} &= 0 \\
a_{41} + a_{32} + a_{23} - a_{11} - a_{12} - a_{13} &= 0.
\end{aligned}
$$

In general, the equation involving the descending diagonal is

$$a_{22} + a_{33} + \cdots + a_{nn} - a_{12} - a_{13} - \cdots - a_{1n} = 0 \qquad (r)$$

and the equation involving the ascending diagonal is

$$a_{n1} + a_{n-12} + \cdots + a_{2n-1} - a_{11} - a_{12} - \cdots - a_{1n-1} = 0. \qquad (c)$$

Prove that if $n \geq 3$, then the $2n$ equations asserting that a matrix is a generalized magic square are linearly independent.

Hint. Equations are really linear forms, so find some matrix annihilated by all equations except equation r, and some matrix annihilated by all equations except equation c.

Problem B2 (30 pts). Let A be an $m \times n$ matrix and B be an $n \times m$ matrix.

(1) Prove that

$$\det(I_m - AB) = \det(I_n - BA).$$

Hint. Consider the matrices

$$X = \begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} I_m & 0 \\ -B & I_n \end{pmatrix}.$$

(2) Prove that

$$\lambda^n \det(\lambda I_m - AB) = \lambda^m \det(\lambda I_n - BA).$$

Hint. Consider the matrices

$$X = \begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} I_m & 0 \\ -B & \lambda I_n \end{pmatrix}.$$

Problem B3 (70 pts). Given a field K (say $K = \mathbb{R}$ or $K = \mathbb{C}$), given any two polynomials $p(X), q(X) \in K[X]$, we says that $q(X)$ *divides* $p(X)$ (and that $p(X)$ *is a multiple of* $q(X)$) iff there is some polynomial $s(X) \in K[X]$ such that

$$p(X) = q(X)s(X).$$

In this case we say that $q(X)$ *is a factor of* $p(X)$, and if $q(X)$ has degree at least one, we say that $q(X)$ *is a nontrivial factor of* $p(X)$.

Let $f(X)$ and $g(X)$ be two polynomials in $K[X]$ with

$$f(X) = a_0X^m + a_1X^{m-1} + \cdots + a_m$$

of degree $m \geq 1$ and

$$g(X) = b_0X^n + b_1X^{n-1} + \cdots + b_n$$

of degree $n \geq 1$ (with $a_0, b_0 \neq 0$).

You will need the following result which you need not prove:

Two polynomials $f(X)$ and $g(X)$ with $\deg(f) = m \geq 1$ *and* $\deg(g) = n \geq 1$ *have some common nontrivial factor iff there exist two nonzero polynomials $p(X)$ and $q(X)$ such that*

$$fp = gq,$$

with $\deg(p) \leq n - 1$ *and* $\deg(q) \leq m - 1$.

(1) Let $\mathcal{P}_m$ denote the vector space of all polynomials in $K[X]$ of degree at most $m-1$, and let $T\colon \mathcal{P}_n \times \mathcal{P}_m \to \mathcal{P}_{m+n}$ be the map given by

$$T(p,q) = fp + gq, \quad p \in \mathcal{P}_n,\ q \in \mathcal{P}_m,$$

where f and g are some fixed polynomials of degree $m \geq 1$ and $n \geq 1$.

Prove that the map T is linear.

(2) Prove that T is not injective iff f and g have a common nontrivial factor.

(3) Prove that f and g have a nontrivial common factor iff $R(f,g) = 0$, where $R(f,g)$ is the determinant given by

$$R(f,g) = \begin{vmatrix}
a_0 & a_1 & \cdots & \cdots & a_m & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & a_0 & a_1 & \cdots & \cdots & a_m & 0 & \cdots & \cdots & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & a_0 & a_1 & \cdots & \cdots & a_m \\
b_0 & b_1 & \cdots & \cdots & \cdots & \cdots & \cdots & b_n & 0 & \cdots & 0 \\
0 & b_0 & b_1 & \cdots & \cdots & \cdots & \cdots & \cdots & b_n & 0 & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & b_0 & b_1 & \cdots & \cdots & \cdots & \cdots & \cdots & b_n
\end{vmatrix}.$$

The above determinant is called the *resultant of f and g*.

Note that the matrix of the resultant is an $(n+m)\times(n+m)$ matrix, with the first row (involving the a_is) occuring n times, each time shifted over to the right by one column, and the $(n+1)$th row (involving the b_js) occuring m times, each time shifted over to the right by one column.

Hint. Express the matrix of T over some suitable basis.

(4) Compute the resultant in the following three cases:

(a) $m = n = 1$, and write $f(X) = aX + b$ and $g(X) = cX + d$.

(b) $m = 1$ and $n \geq 2$ arbitrary.

(c) $f(X) = aX^2 + bX + c$ and $g(X) = 2aX + b$.

Extra Credit (40 pts). Compute the resultant of $f(X) = X^3 + pX + q$ and $g(X) = 3X^2 + p$, and

$$\begin{aligned} f(X) &= a_0X^2 + a_1X + a_2 \\ g(X) &= b_0X^2 + b_1X + b_2. \end{aligned}$$

In the second case, you should get

$$4R(f,g) = (2a_0b_2 - a_1b_1 + 2a_2b_0)^2 - (4a_0a_2 - a_1^2)(4b_0b_2 - b_1^2).$$

Problem B4 (40 pts). Give an example of a norm on $\mathbb{C}^n$ and of a *real* matrix A such that

$$\|A\|_{\mathbb{R}} < \|A\|,$$

where $\|-\|_{\mathbb{R}}$ and $\|-\|$ are the operator norms associated with the vector norm $\|-\|$, as defined in the notes.

Hint. This can already be done for $n = 2$.

Problem B5 (30 pts). Let $\|\ \|$ be any operator norm. Given an invertible $n \times n$ matrix A, if $c = 1/(2\,\|A^{-1}\|)$, then for every $n \times n$ matrix H, if $\|H\| \leq c$, then $A + H$ is invertible. Furthermore, show that if $\|H\| \leq c$, then $\|(A+H)^{-1}\| \leq 1/c$.

Problem B6 (40 pts). Let A be any $m \times n$ matrix and let $\lambda \in \mathbb{R}$ be any positive real number $\lambda > 0$.

(1) Prove that $A^\top A + \lambda I_n$ and $AA^\top + \lambda I_m$ are invertible.

(2) Prove that

$$A^\top(AA^\top + \lambda I_m)^{-1} = (A^\top A + \lambda I_m)^{-1}A^\top.$$

Remark: The expressions above correspond to the matrix for which the function

$$\Phi(x) = (Ax - b)^\top (Ax - b) + \lambda x^\top x$$

achieves a minimum. It shows up in machine learning (kernel methods).

Problem B7 (80 pts). Let A be a real 2×2 matrix

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} \\ a_{2\,1} & a_{2\,2} \end{pmatrix}.$$

(1) Prove that the squares of the singular values $\sigma_1 \geq \sigma_2$ of A are the roots of the quadratic equation

$$X^2 - \mathrm{tr}(A^\top A)X + |\det(A)|^2 = 0.$$

(2) If we let

$$\mu(A) = \frac{a_{1\,1}^2 + a_{1\,2}^2 + a_{2\,1}^2 + a_{2\,2}^2}{2|a_{1\,1}a_{2\,2} - a_{1\,2}a_{2\,1}|},$$

prove that

$$\mathrm{cond}_2(A) = \frac{\sigma_1}{\sigma_2} = \mu(A) + (\mu(A)^2 - 1)^{1/2}.$$

(3) Consider the subset $\mathcal{S}$ of 2×2 invertible matrices whose entries $a_{i\,j}$ are integers such that $0 \leq a_{ij} \leq 100$.

Prove that the functions $\mathrm{cond}_2(A)$ and $\mu(A)$ reach a maximum on the set $\mathcal{S}$ for the same values of A.

Check that for the matrix

$$A_m = \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix}$$

we have

$$\mu(A_m) = 19,603 \quad \det(A_m) = -1$$

and

$$\mathrm{cond}_2(A_m) \approx 39,206.$$

(4) Prove that for all $A \in \mathcal{S}$, if $|\det(A)| \geq 2$ then $\mu(A) \leq 10,000$. Conclude that the maximum of $\mu(A)$ on $\mathcal{S}$ is achieved for matrices such that $\det(A) = \pm 1$. Prove that finding matrices that maximize μ on $\mathcal{S}$ is equivalent to finding some integers n_1, n_2, n_3, n_4 such that

$$\begin{aligned}
&0 \leq n_4 \leq n_3 \leq n_2 \leq n_1 \leq 100 \\
&n_1^2 + n_2^2 + n_3^2 + n_4^2 \geq 100^2 + 99^2 + 99^2 + 98^2 = 39,206 \\
&|n_1 n_4 - n_2 n_3| = 1.
\end{aligned}$$

You may use without proof that the fact that the only solution to the above constraints is the multiset

$$\{100, 99, 99, 98\}.$$

(5) Deduce from part (4) that the matrices in $\mathcal{S}$ for which μ has a maximum value are

$$A_m = \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix} \quad \begin{pmatrix} 98 & 99 \\ 99 & 100 \end{pmatrix} \quad \begin{pmatrix} 99 & 100 \\ 98 & 99 \end{pmatrix} \quad \begin{pmatrix} 99 & 98 \\ 100 & 99 \end{pmatrix}$$

and check that μ has the same value for these matrices. Conclude that

$$\max_{A \in \mathcal{S}} \mathrm{cond}_2(A) = \mathrm{cond}_2(A_m).$$

(6) Solve the system

$$\begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 199 \\ 197 \end{pmatrix}.$$

Perturb the right-hand side b by

$$\delta b = \begin{pmatrix} -0.0097 \\ 0.0106 \end{pmatrix}$$

and solve the new system

$$A_m y = b + \delta b$$

where $y = (y_1, y_2)$. Check that

$$\delta x = y - x = \begin{pmatrix} 2 \\ -2.0203 \end{pmatrix}.$$

Compute $\|x\|_2$, $\|\delta x\|_2$, $\|b\|_2$, $\|\delta b\|_2$, and estimate

$$c = \frac{\|\delta x\|_2}{\|x\|_2} \left(\frac{\|\delta b\|_2}{\|b\|_2} \right)^{-1}.$$

Ckeck that

$$c \approx \mathrm{cond}_2(A_m) = 39,206.$$

Problem B8 (20 pts). Let $\varphi \colon E \times E \to \mathbb{R}$ be a bilinear form on a real vector space E of finite dimension n. Given any basis $(e_1, \ldots, e_n)$ of E, let $A = (a_{i\,j})$ be the matrix defined such that

$$a_{i\,j} = \varphi(e_i, e_j),$$

$1 \leq i, j \leq n$. We call A *the matrix of* φ *w.r.t. the basis* $(e_1, \ldots, e_n)$.

(a) For any two vectors x and y, if X and Y denote the column vectors of coordinates of x and y w.r.t. the basis $(e_1, \ldots, e_n)$, prove that

$$\varphi(x, y) = X^\top AY.$$

(b) Recall that A is a *symmetric* matrix if $A = A^\top$. Prove that φ is symmetric if A is a symmetric matrix.

(c) If $(f_1, \ldots, f_n)$ is another basis of E and P is the change of basis matrix from $(e_1, \ldots, e_n)$ to $(f_1, \ldots, f_n)$, prove that the matrix of φ w.r.t. the basis $(f_1, \ldots, f_n)$ is

$$P^\top AP.$$

The common rank of all matrices representing φ is called the *rank* of φ.

Problem B9 (40 pts). Let $\varphi \colon E \times E \to \mathbb{R}$ be a symmetric bilinear form on a real vector space E of finite dimension n. Two vectors x and y are said to be *conjugate or orthogonal w.r.t.* φ if $\varphi(x, y) = 0$. The main purpose of this problem is to prove that there is a basis of vectors that are pairwise conjugate w.r.t. φ.

(a) Prove that if $\varphi(x, x) = 0$ for all $x \in E$, then φ is identically null on E.

Otherwise, we can assume that there is some vector $x \in E$ such that $\varphi(x, x) \neq 0$.

Use induction to prove that there is a basis of vectors $(u_1, \ldots, u_n)$ that are pairwise conjugate w.r.t. φ.

Hint. For the induction step, proceed as follows. Let $(u_1, e_2, \ldots, e_n)$ be a basis of E, with $\varphi(u_1, u_1) \neq 0$. Prove that there are scalars $\lambda_2, \ldots, \lambda_n$ such that each of the vectors

$$v_i = e_i + \lambda_i u_1$$

is conjugate to u_1 w.r.t. φ, where $2 \leq i \leq n$, and that $(u_1, v_2, \ldots, v_n)$ is a basis.

(b) Let $(e_1, \ldots, e_n)$ be a basis of vectors that are pairwise conjugate w.r.t. φ, and assume that they are ordered such that

$$\varphi(e_i, e_i) = \begin{cases} \theta_i \neq 0 & \text{if } 1 \leq i \leq r, \\ 0 & \text{if } r + 1 \leq i \leq n, \end{cases}$$

where r is the rank of φ. Show that the matrix of φ w.r.t. $(e_1, \ldots, e_n)$ is a diagonal matrix, and that

$$\varphi(x, y) = \sum_{i=1}^{r} \theta_i x_i y_i,$$

where $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{i=1}^n y_i e_i$.

Prove that for every symmetric matrix A, there is an invertible matrix P such that

$$P^\top AP = D,$$

where D is a diagonal matrix.

(c) Prove that there is an integer p, $0 \leq p \leq r$ (where r is the rank of φ), such that $\varphi(u_i, u_i) > 0$ for exactly p vectors of every basis $(u_1, \ldots, u_n)$ of vectors that are pairwise conjugate w.r.t. φ (*Sylvester's inertia theorem*).

Proceed as follows. Assume that in the basis $(u_1, \ldots, u_n)$, for any $x \in E$, we have

$$\varphi(x, x) = \alpha_1 x_1^2 + \cdots + \alpha_p x_p^2 - \alpha_{p+1} x_{p+1}^2 - \cdots - \alpha_r x_r^2,$$

where $x = \sum_{i=1}^n x_i u_i$, and that in the basis $(v_1, \ldots, v_n)$, for any $x \in E$, we have

$$\varphi(x, x) = \beta_1 y_1^2 + \cdots + \beta_q y_q^2 - \beta_{q+1} y_{q+1}^2 - \cdots - \beta_r y_r^2,$$

where $x = \sum_{i=1}^n y_i v_i$, with $\alpha_i > 0$, $\beta_i > 0$, $1 \leq i \leq r$.

Assume that $p > q$ and derive a contradiction. First, consider x in the subspace F spanned by

$$(u_1, \ldots, u_p, u_{r+1}, \ldots, u_n),$$

and observe that $\varphi(x, x) \geq 0$ if $x \neq 0$. Next, consider x in the subspace G spanned by

$$(v_{q+1}, \ldots, v_r),$$

and observe that $\varphi(x, x) < 0$ if $x \neq 0$. Prove that $F \cap G$ is nontrivial (i.e., contains some nonnull vector), and derive a contradiction. This implies that $p \leq q$. Finish the proof.

The pair $(p, r - p)$ is called the *signature* of φ.

(d) A symmetric bilinear form φ is *definite* if for every $x \in E$, if $\varphi(x, x) = 0$, then $x = 0$.

Prove that a symmetric bilinear form is definite iff its signature is either $(n, 0)$ or $(0, n)$. In other words, a symmetric definite bilinear form has rank n and is either positive or negative.

TOTAL: 400 + 40 points.