Fall, 2015 CIS 515

Fundamentals of Linear Algebra and Optimization
Jean Gallier

Homework 5

November 9, 2015; Due November 24 2015
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 (80 pts). (1) Implement the method for converting a rectangular matrix to reduced row echelon fom.

(2) Use the above method to find the inverse of an invertible $n \times n$ matrix A, by applying it to the the $n \times 2n$ matrix $[A\, I]$ obtained by adding the n columns of the identity matrix to A.

(3) Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ 2 & 3 & 4 & 5 & \cdots & n+1 \\ 3 & 4 & 5 & 6 & \cdots & n+2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ n & n+1 & n+2 & n+3 & \cdots & 2n-1 \end{pmatrix}.$$

Using your program, find the row reduced echelon form of A for $n = 4, \ldots, 20$.

Also run the `Matlab` `rref` function and compare results.

Your program probably disagrees with `rref` even for small values of n. The problem is that some pivots are very small and the normalization step (to make the pivot 1) causes roundoff errors. Use a tolerance parameter to fix this problem.

What can you conjecture about the rank of A?

(4) Prove that the matrix A has the following row reduced form:

$$R = \begin{pmatrix} 1 & 0 & -1 & -2 & \cdots & -(n-2) \\ 0 & 1 & 2 & 3 & \cdots & n-1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Deduce from the above that A has rank 2.

Hint. Some well chosen sequence of row operations.

(5) Use your program to show that if you add any number greater than or equal to $(2/25)n^2$ to every diagonal entry of A you get an invertible matrix! In fact, running the `Matlab` fuction `chol` should tell you that these matrices are SPD (symmetric, positive definite).

Remark: The above phenomenon will be explained in Problem B4. If you have a rigorous and *simple* explanation for this phenomenon, let me know!

Problem B4 (120 pts). The purpose of this problem is to prove that the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ 2 & 3 & 4 & 5 & \cdots & n+1 \\ 3 & 4 & 5 & 6 & \cdots & n+2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ n & n+1 & n+2 & n+3 & \cdots & 2n-1 \end{pmatrix}$$

is

$$P_A(\lambda) = \lambda^{n-2}\left(\lambda^2 - n^2\lambda - \frac{1}{12}n^2(n^2-1)\right).$$

(1) Prove that the characteristic polynomial $P_A(\lambda)$ is given by

$$P_A(\lambda) = \lambda^{n-2} P(\lambda),$$

with

$$P(\lambda) = \begin{vmatrix} \lambda-1 & -2 & -3 & -4 & \cdots & -n+3 & -n+2 & -n+1 & -n \\ -\lambda-1 & \lambda-1 & -1 & -1 & \cdots & -1 & -1 & -1 & -1 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{vmatrix}.$$

(2) Prove that the sum of the roots λ_1, λ_2 of the (degree two) polynomial $P(\lambda)$ is

$$\lambda_1 + \lambda_2 = n^2.$$

The problem is thus to compute the product $\lambda_1\lambda_2$ of these roots. Prove that

$$\lambda_1\lambda_2 = P(0).$$

(3) The problem is now to evaluate $d_n = P(0)$, where

$$d_n = \begin{vmatrix} -1 & -2 & -3 & -4 & \cdots & -n+3 & -n+2 & -n+1 & -n \\ -1 & -1 & -1 & -1 & \cdots & -1 & -1 & -1 & -1 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{vmatrix}$$

I suggest the following strategy: cancel out the first entry in row 1 and row 2 by adding a suitable multiple of row 3 to row 1 and row 2, and then subtract row 2 from row 1. Expand the determinant according to the first column.

You will notice that the first two entries on row 1 and the first two entries on row 2 change, but the rest of the matrix looks the same, except that the dimension is reduced.

This suggests setting up a recurrence involving the entries u_k, v_k, x_k, y_k in the determinant

$$D_k = \begin{vmatrix} u_k & x_k & -3 & -4 & \cdots & -n+k-3 & -n+k-2 & -n+k-1 & -n+k \\ v_k & y_k & -1 & -1 & \cdots & -1 & -1 & -1 & -1 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{vmatrix},$$

starting with $k = 0$, with

$$u_0 = -1, \quad v_0 = -1, \quad x_0 = -2, \quad y_0 = -1,$$

and ending with $k = n-2$, so that

$$d_n = D_{n-2} = \begin{vmatrix} u_{n-3} & x_{n-3} & -3 \\ v_{n-3} & y_{n-3} & -1 \\ 1 & -2 & 1 \end{vmatrix} = \begin{vmatrix} u_{n-2} & x_{n-2} \\ v_{n-2} & y_{n-2} \end{vmatrix}.$$

Prove that we have the recurrence relations

$$\begin{pmatrix} u_{k+1} \\ v_{k+1} \\ x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 & -1 \\ 0 & 2 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_k \\ v_k \\ x_k \\ y_k \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -2 \\ -1 \end{pmatrix}.$$

These appear to be nasty affine recurrence relations, so we will use the trick to convert this affine map to a linear map.

(4) Consider the linear map given by

$$\begin{pmatrix} u_{k+1} \\ v_{k+1} \\ x_{k+1} \\ y_{k+1} \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 & -1 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & -2 \\ 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_k \\ v_k \\ x_k \\ y_k \\ 1 \end{pmatrix},$$

and show that its action on u_k, v_k, x_k, y_k is the same as the affine action of part (3).

Use `Matlab` to find the eigenvalues of the matrix

$$T = \begin{pmatrix} 2 & -2 & 1 & -1 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & -2 \\ 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

You will be stunned!

Let N be the matrix given by

$$N = T - I.$$

Prove that

$$N^4 = 0.$$

Use this to prove that

$$T^k = I + kN + \frac{1}{2}k(k-1)N^2 + \frac{1}{6}k(k-1)(k-2)N^3,$$

for all $k \geq 0$.

(5) Prove that

$$\begin{pmatrix} u_k \\ v_k \\ x_k \\ y_k \\ 1 \end{pmatrix} = T^k \begin{pmatrix} -1 \\ -1 \\ -2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 & -1 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & -2 \\ 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^k \begin{pmatrix} -1 \\ -1 \\ -2 \\ -1 \\ 1 \end{pmatrix},$$

for $k \geq 0$.

Prove that

$$T^k = \begin{pmatrix} k+1 & -k(k+1) & k & -k^2 & \frac{1}{6}(k-1)k(2k-7) \\ 0 & k+1 & 0 & k & -\frac{1}{2}(k-1)k \\ -k & k^2 & 1-k & (k-1)k & -\frac{1}{3}k((k-6)k+11) \\ 0 & -k & 0 & 1-k & \frac{1}{2}(k-3)k \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and thus, that

$$\begin{pmatrix} u_k \\ v_k \\ x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \frac{1}{6}(2k^3+3k^2-5k-6) \\ -\frac{1}{2}(k^2+3k+2) \\ \frac{1}{3}(-k^3+k-6) \\ \frac{1}{2}(k^2+k-2) \end{pmatrix},$$

and that

$$\begin{vmatrix} u_k & x_k \\ v_k & y_k \end{vmatrix} = -1 - \frac{7}{3}k - \frac{23}{12}k^2 - \frac{2}{3}k^3 - \frac{1}{12}k^4.$$

As a consequence, prove that amazingly,

$$d_n = D_{n-2} = -\frac{1}{12}n^2(n^2-1).$$

(6) Prove that the characteristic polynomial of A is indeed

$$P_A(\lambda) = \lambda^{n-2}\left(\lambda^2 - n^2\lambda - \frac{1}{12}n^2(n^2-1)\right).$$

Use the above to show that the two nonzero eigenvalues of A are

$$\lambda = \frac{n}{2}\left(n \pm \frac{\sqrt{3}}{3}\sqrt{4n^2-1}\right).$$

The negative eigenvalue λ_1 can also be expressed as

$$\lambda_1 = n^2\frac{(3-2\sqrt{3})}{6}\sqrt{1-\frac{1}{4n^2}}.$$

Use this expression to explain the phenomenon in B3(5): If we add any number greater than or equal to $(2/25)n^2$ to every diagonal entry of A we get an invertible matrix. What about $0.077351n^2$? Try it!

Extra Credit Problem B5 (100 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 B6 (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 B7 (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

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

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 B8 (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 $\text{cond}_2(A)$ 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

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

(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

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

(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

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

TOTAL: 440 + 100 points.