18.06 Problem Set 6 Solutions

Due Thursday, 18 March 2010 at 4 pm in 2-108.
Total: 100 points

Section 4.3. Problem 4: Write down $E = \|Ax - b\|^2$ as a sum of four squares– the last one is $(C + 4D - 20)^2$. Find the derivative equations $\partial E/\partial C = 0$ and $\partial E/\partial D = 0$. Divide by 2 to obtain the normal equations $A^T A\widehat{x} = A^T b$.

Solution (4 points)

Observe

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}, \text{ and define } x = \begin{pmatrix} C \\ D \end{pmatrix}.$$

Then

$$Ax - b = \begin{pmatrix} C \\ C + D - 8 \\ C + 3D - 8 \\ C + 4D - 20 \end{pmatrix},$$

and

$$\|Ax - b\|^2 = C^2 + (C + D - 8)^2 + (C + 3D - 8)^2 + (C + 4D - 20)^2.$$

The partial derivatives are

$$\partial E/\partial C = 2C + 2(C + D - 8) + 2(C + 3D - 8) + 2(C + 4D - 20) = 8C + 16D - 72,$$

$$\partial E/\partial D = 2(C + D - 8) + 6(C + 3D - 8) + 8(C + 4D - 20) = 16C + 52D - 224.$$

On the other hand,

$$A^T A = \begin{pmatrix} 4 & 8 \\ 8 & 26 \end{pmatrix}, \; A^T b = \begin{pmatrix} 36 \\ 112 \end{pmatrix}.$$

Thus, $A^T Ax = A^T b$ yields the equations $4C + 8D = 36$, $8C + 26D = 112$. Multiplying by 2 and looking back, we see that these are precisely the equations $\partial E/\partial C = 0$ and $\partial E/\partial D = 0$.

Section 4.3. Problem 7: Find the closest line $b = Dt$, *through the origin*, to the same four points. An exact fit would solve $D \cdot 0 = 0$, $D \cdot 1 = 8$, $D \cdot 3 = 8$, $D \cdot 4 = 20$.

Find the 4 by 1 matrix A and solve $A^T A\widehat{x} = A^T b$. Redraw figure 4.9a showing the best line $b = Dt$ and the e's.

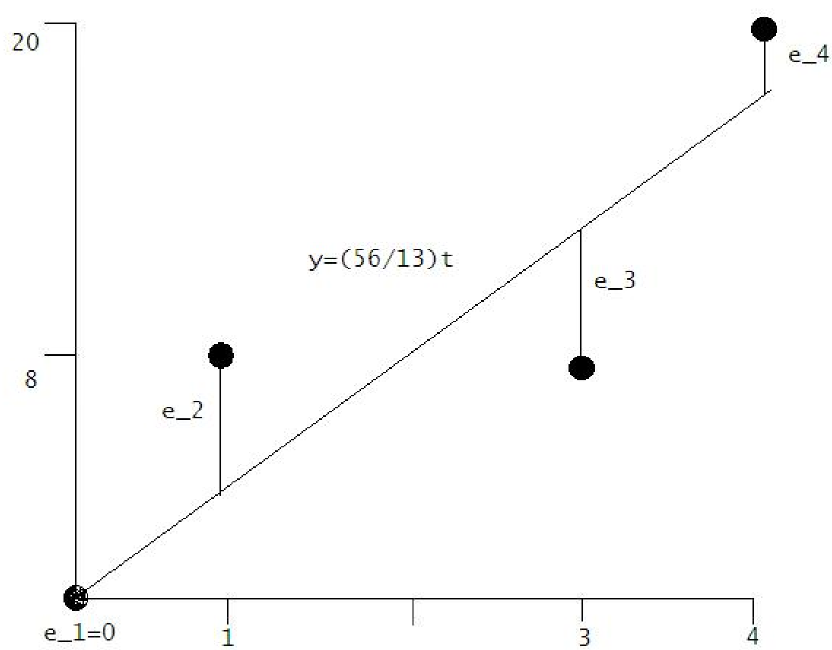
Solution (4 points) Observe

$$A = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 4 \end{pmatrix}, \ b = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}, \ A^T A = (26), \ A^T b = (112).$$

Thus, solving $A^T Ax = A^T b$, we arrive at

$$D = 56/13.$$

Here is the diagram analogous to figure 4.9a.



Section 4.3. Problem 9: Form the closest parabola $b = C + Dt + Et^2$ to the same four points, and write down the unsolvable equations $Ax = b$ in three unknowns

$x = (C, D, E)$. Set up the three normal equations $A^T A\widehat{x} = A^T b$ (solution not required). In figure 4.9a you are now fitting a parabola to 4 points–what is happening in Figure 4.9b?

Solution (4 points)
Note

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix},\ b = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix},\ x = \begin{pmatrix} C \\ D \\ E \end{pmatrix}.$$

Then multiplying out $Ax = b$ yields the equations

$$C = 0,\ C + D + E = 8,\ C + 3D + 9E = 8,\ C + 4D + 16E = 20.$$

Take the sum of the fourth equation and twice the second equation and subtract the sum of the first equation and two times the third equation. One gets $0 = 20$. Hence, these equations are not simultaneously solvable.
Computing, we get

$$A^T A = \begin{pmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{pmatrix},\ A^T b = \begin{pmatrix} 36 \\ 112 \\ 400 \end{pmatrix}.$$

Thus, solving this problem is the same as solving the system

$$\begin{pmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 36 \\ 112 \\ 400 \end{pmatrix}.$$

The analogue of diagram 4.9(b) in this case would show three vectors $a_1 = (1, 1, 1, 1)$, $a_2 = (0, 1, 3, 4)$, $a_3 = (0, 1, 9, 16)$ spanning a three dimensional vector subspace of $\mathbb{R}^4$. It would also show the vector $b = (0, 8, 8, 20)$, and the projection $p = Ca_1 + Da_2 + Ea_3$ of b into the three dimensional subspace.

Section 4.3. Problem 26: Find the *plane* that gives the best fit to the 4 values $b = (0, 1, 3, 4)$ at the corners $(1, 0)$ and $(0, 1)$ and $(-1, 0)$ and $(0, -1)$ of a square. The equations $C + Dx + Ey = b$ at those 4 points are $Ax = b$ with 3 unknowns $x = (C, D, E)$. What is A? At the center $(0, 0)$ of the square, show that $C + Dx + Ey$ is the average of the b's.

Solution (12 points)

Note

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

To find the best fit plane, we must find x such that $Ax - b$ is in the left nullspace of A. Observe

$$Ax - b = \begin{pmatrix} C + D \\ C + E - 1 \\ C - D - 3 \\ C - E - 4 \end{pmatrix}.$$

Computing, we find that the first entry of $A^T(Ax - b)$ is $4C - 8$. This is zero when $C = 2$, the average of the entries of b. Plugging in the point $(0, 0)$, we get $C + D(0) + E(0) = C = 2$ as desired.

Section 4.3. Problem 29: Usually there will be exactly one hyperplane in $\mathbb{R}^n$ that contains the n given points $x = 0, a_1, \ldots, a_{n-1}$. (Example for n=3: There will be exactly one plane containing $0, a_1, a_2$ unless ________.) What is the test to have exactly one hyperplane in $\mathbb{R}^n$?

Solution (12 points)

The sentence in paranthesis can be completed a couple of different ways. One could write "There will be exactly one plane containing $0, a_1, a_2$ unless these three points are colinear". Another acceptable answer is "There will be exactly one plane containing $0, a_1, a_2$ unless the vectors a_1 and a_2 are linearly dependent".

In general, $0, a_1, \ldots, a_{n-1}$ will be contained in an unique hyperplane unless all of the points $0, a_1, \ldots, a_{n-1}$ are contained in an $n - 2$ dimensional subspace. Said another way, $0, a_1, \ldots, a_{n-1}$ will be contained in an unique hyperplane unless the vectors $a_1, \ldots, a_{n-1}$ are linearly dependent.

Section 4.4. Problem 10: Orthonormal vectors are automatically linearly independent.
(a) Vector proof: When $c_1q_1 + c_2q_2 + c_3q_3 = 0$, what dot product leads to $c_1 = 0$? Similarly $c_2 = 0$ and $c_3 = 0$. Thus, the q's are independent.
(b) Matrix proof: Show that $Qx = 0$ leads to $x = 0$. Since Q may be rectangular, you can use Q^T but not Q^{-1}.

Solution (4 points) For part (a): Dotting the expression $c_1q_1 + c_2q_2 + c_3q_3$ with q_1, we get $c_1 = 0$ since $q_1 \cdot q_1 = 1$, $q_1 \cdot q_2 = q_1 \cdot q_3 = 0$. Similarly, dotting the expression with q_2 yields $c_2 = 0$ and dotting the expression with q_3 yields $c_3 = 0$. Thus, $\{q_1, q_2, q_3\}$ is a linearly independent set.

For part (b): Let Q be the matrix whose columns are q_1, q_2, q_3. Since Q has orthonormal columns, Q^TQ is the three by three identity matrix. Now, multiplying the equation $Qx = 0$ on the left by Q^T yields $x = 0$. Thus, the nullspace of Q is the zero vector and its columns are linearly independent.

Section 4.4. Problem 18: Find the orthonormal vectors A, B, C by Gram-Schmidt from a, b, c:

$$a = (1, -1, 0, 0) \quad b = (0, 1, -1, 0) \quad c = (0, 0, 1, -1).$$

Show $\{A, B, C\}$ and $\{a, b, c\}$ are bases for the space of vectors perpendicular to $d = (1, 1, 1, 1)$.

Solution (4 points) We apply Gram-Schmidt to a, b, c. We have

$$A = \frac{a}{\|a\|} = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0, 0\right).$$

Next,

$$B = \frac{b - (b \cdot A)A}{\|b - (b \cdot A)A\|} = \frac{(\frac{1}{2}, \frac{1}{2}, -1, 0)}{\|(\frac{1}{2}, \frac{1}{2}, -1, 0)\|} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\sqrt{\frac{2}{3}}, 0\right).$$

Finally,

$$C = \frac{c - (c \cdot A)A - (c \cdot B)B}{\|c - (c \cdot A)A - (c \cdot B)B\|} = \left(\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, -\frac{\sqrt{3}}{2}\right).$$

Note that $\{a, b, c\}$ is a linearly independent set. Indeed,

$$x_1a + x_2b + x_3c = (x_1, x_2 - x_1, x_3 - x_2, -x_3) = (0, 0, 0, 0)$$

implies that $x_1 = x_2 = x_3 = 0$. We check $a \cdot (1, 1, 1, 1) = b \cdot (1, 1, 1, 1) = c \cdot (1, 1, 1, 1) = 0$. Hence, all three vectors are in the nullspace of $(1, 1, 1, 1)$. Moreover, the dimension of the column space of the transpose and the dimension of the nullspace sum to the dimension of $\mathbb{R}^4$. Thus, the space of vectors perpendicular to $(1, 1, 1, 1)$ is three dimensional. Since $\{a, b, c\}$ is a linearly independent set in this space, it is a basis.

Since $\{A, B, C\}$ is an orthonormal set, it is a linearly independent set by problem 10. Thus, it must also span the space of vectors perpendicular to $(1, 1, 1, 1)$, and it is also a basis of this space.

Section 4.4. Problem 35: Factor $[Q, R] = \mathbf{qr}(\mathbf{A})$ for $A = \mathbf{eye}(\mathbf{4}) - \mathbf{diag}([\mathbf{111}], -\mathbf{1})$. You are orthogonalizing the columns $(1, -1, 0, 0)$, $(0, 1, -1, 0)$, $(0, 0, 1, -1)$, and $(0, 0, 0, 1)$ of A. Can you scale the orthogonal columns of Q to get nice integer components?

Solution (12 points) Here is a copy of the matlab code

```
>> A=eye(4)-diag([1 1 1],-1)

A =

     1     0     0     0
    -1     1     0     0
     0    -1     1     0
     0     0    -1     1

>> [Q,R]=qr(A)

Q =

   -0.7071   -0.4082   -0.2887    0.5000
    0.7071   -0.4082   -0.2887    0.5000
         0    0.8165   -0.2887    0.5000
         0         0    0.8660    0.5000


R =

   -1.4142    0.7071         0         0
         0   -1.2247    0.8165         0
         0         0   -1.1547    0.8660
         0         0         0    0.5000
```

Note that scaling the first column by $\sqrt{2}$, the second column by $\sqrt{6}$, the third column by $2\sqrt{3}$, and the fourth column by 2 yields

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & -3 & 1 \end{pmatrix}.$$

Section 4.4. Problem 36: If A is m by n, $\mathbf{qr(A)}$ produces a *square* A and zeroes below R: The factors from MATLAB are $(m \text{ by } m)(m \text{ by } n)$

$$A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}.$$

The n columns of Q_1 are an orthonormal basis for which fundamental subspace? The $m-n$ columns of Q_2 are an orthonormal basis for which fundamental subspace?

Solution (12 points) The n columns of Q_1 form an orthonormal basis for the column space of A. The $m-n$ columns of Q_2 form an orthonormal basis for the left nullspace of A.

Section 5.1. Problem 10: If the entries in every row of A add to zero, solve $Ax = 0$ to prove $\det A = 0$. If those entries add to one, show that $\det(A - I) = 0$. Does this mean $\det A = I$?

Solution (4 points) If $x = (1, 1, \ldots, 1)$, then the components of Ax are the sums of the rows of A. Thus, $Ax = 0$. Since A has non-zero nullspace, it is not invertible and $\det A = 0$. If the entries in every row of A sum to one, then the entries in every row of $A-I$ sum to zero. Hence, $A-I$ has a non-zero nullspace and $\det(A-I) = 0$. This does not mean that $\det A = I$. For example if

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

then the entries of every row of A sum to one. However, $\det A = -1$.

Section 5.1. Problem 18: Use row operations to show that the 3 by 3 "Vandermonde determinant" is

$$\det \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} = (b-a)(c-a)(c-b).$$

Solution (4 points) Doing elimination, we get

$$\det \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix} = \det \begin{pmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{pmatrix} = (b-a)\det \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & c-a & c^2-a^2 \end{pmatrix} =$$

$$= (b-a)\det\begin{pmatrix}1 & a & a^2\\ 0 & 1 & b+a\\ 0 & 0 & (c-a)(c-b)\end{pmatrix} = (b-a)(c-a)(c-b).$$

Section 5.1. Problem 31: (MATLAB) The Hilbert matrix **hilb(n)** has i, j entry equal to $1/(i+j-1)$. Print the determinants of $\mathbf{hilb(1)}, \mathbf{hilb(2)}, \ldots, \mathbf{hilb(10)}$. Hilbert matrices are hard to work with! What are the pivots of **hilb(5)**?

Solution (12 points) Here is the relevant matlab code.

```
>> [det(hilb(1)) det(hilb(2)) det(hilb(3)) det(hilb(4))
det(hilb(5)) det(hilb(6)) det(hilb(7)) det(hilb(8))
det(hilb(9)) det(hilb(10))]

ans =

    1.0000    0.0833    0.0005    0.0000    0.0000
    0.0000    0.0000    0.0000    0.0000    0.0000

>> [L,U,P]=lu(hilb(5))

L =

    1.0000         0         0         0         0
    0.3333    1.0000         0         0         0
    0.5000    1.0000    1.0000         0         0
    0.2000    0.8000   -0.9143    1.0000         0
    0.2500    0.9000   -0.6000    0.5000    1.0000


U =

    1.0000    0.5000    0.3333    0.2500    0.2000
         0    0.0833    0.0889    0.0833    0.0762
         0         0   -0.0056   -0.0083   -0.0095
         0         0         0    0.0007    0.0015
         0         0         0         0   -0.0000


P =

     1     0     0     0     0
     0     0     1     0     0
     0     1     0     0     0
     0     0     0     0     1
     0     0     0     1     0
```

Note that the determinants of the 4th through 10th Hilbert matrices differ from zero by less than one ten thousandth. The pivots of the fifth Hilbert matrix are $1, .0833, -.0056, .0007, .0000$ up to four significant figures. Thus, we see that there is even a pivot of the fifth Hilbert matrix that differs from zero by less than one ten thousandth.

Section 5.1. Problem 32: (MATLAB) What is the typical determinant (experimentally) of **rand(n)** and **randn(n)** for $n = 50, 100, 200, 400$? (And what does "Inf" mean in MATLAB?)

Solution (12 points) This matlab code computes some examples for rand.

```
>> [det(rand(50)) det(rand(50)) det(rand(50)) det(rand(50))
det(rand(50)) det(rand(50))]

ans =

   1.0e+06 *

   -0.5840   -1.1620   -0.0612    0.3953    0.5149   -0.0436

>> [det(rand(100)) det(rand(100)) det(rand(100)) det(rand(100))
det(rand(100)) det(rand(100))]

ans =

   1.0e+26 *

   -0.6288   -0.0001   -0.1463    0.6322    3.5820    0.0929

>> [det(rand(200)) det(rand(200)) det(rand(200)) det(rand(200))
det(rand(200)) det(rand(200))]

ans =

   1.0e+80 *

   -1.2212    0.0246    0.1505    0.0791    8.4722   -4.5166

>> [det(rand(400)) det(rand(400)) det(rand(400)) det(rand(400))
det(rand(400)) det(rand(400))]

ans =

  1.0e+219 *

    0.4479    1.0835    1.8087    5.5787   -0.3650    5.6855
```

As you can see, **rand(50)** is around 10^5, **rand(100)** is around 10^{25}, **rand(200)** is around 10^{79}, and **rand(400)** is around 10^{219}.
This matlab code computes some examples for randn.

```
>> [det(randn(50)) det(randn(50)) det(randn(50)) det(randn(50))
det(randn(50)) det(randn(50))]

ans =

   1.0e+31 *

    1.2894   -0.0421    0.6148   -0.4418    3.0691   -9.5823

>> [det(randn(100)) det(randn(100)) det(randn(100))
det(randn(100)) det(randn(100)) det(randn(100))]

ans =

   1.0e+78 *

   -0.6426    2.7239   -0.6567    2.1435    1.3960   -1.1224

>> [det(randn(200)) det(randn(200)) det(randn(200))
det(randn(200)) det(randn(200)) det(randn(200))]

ans =

  1.0e+187 *

    1.0414    0.0137    0.1884    0.3810   -0.2961   -1.1438

>> [det(randn(400)) det(randn(400)) det(randn(400))
det(randn(400)) det(randn(400)) det(randn(400))]

ans =

   Inf   Inf  -Inf  -Inf   Inf  -Inf
```

Note that **randn(50)** is around 10^{31}, **randn(100)** is around 10^{78}, **randn(200)** is around 10^{186}, and **randn(400)** is just too big for matlab.