Homework 6
Solutions

Least Squares

1. (Heath 3.5-adapted) Let $\mathbf{x}$ be the solution to the linear least squares problem $A\mathbf{x} \approx \mathbf{b}$, where

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

Let $\mathbf{r} = \mathbf{b} - A\mathbf{x}$ be the corresponding residual vector. Which of the following three vectors is a possible value for $\mathbf{r}$? Why?

(a) $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$

Solution:

The residual vector must satisfy the condition $A^T\mathbf{r} = \mathbf{0}$. The only one of these which satisfies the condition is (b).

2. Let the vectors $\mathbf{x}$ and $\mathbf{y}$ be given by the following Matlab code:

```
x = -5:5;
rng(12);
y = .2 * x .^2  + rand(size(x));
```

Find the best fitting parabola $ax^2 + bx + c$ to the data by solving the least squares problem in Matlab/Octave or another program of your choice. What are the coefficients a, b, c that you found? Include your code.

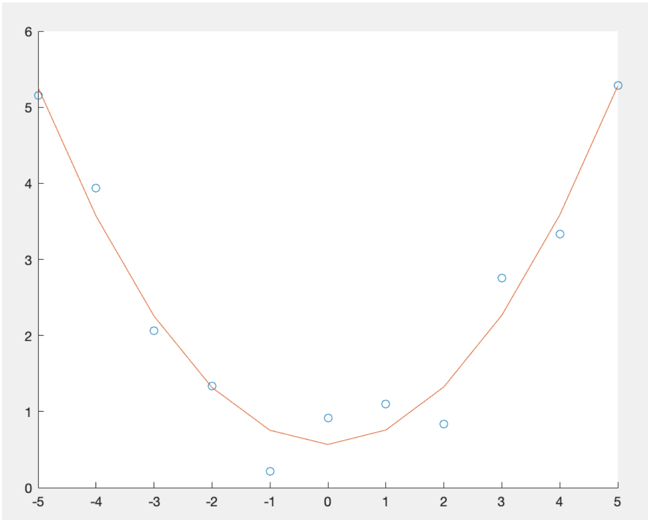
Solution:

Here is some Matlab code to solve the problem and generate a plot of the result (not required).

```
x = -5:5;
rng(12);
y = .2 * x .^2  + rand(size(x));

% construct the matrix A with columns (1, x, x^2)
clear A
A(:,1) = ones(size(x'));
A(:,2) = (x');
A(:,3) = (x') .^ 2;

% construct the rhs
b = y';

% solve for coefficients a = [a0, a1, a2] such that
% y = a0 + a1  x + a2 x^2
% is the bit fit parabola to the data in the ls sense
```

```
a = A \ b

figure
hold on
scatter(x,y)
plot(x,a(1) + a(2) * x + a(3) * x.^2)
```

It gives the solution vector $a = 0.567846630671078, b = 0.001848845233161, c = 0.188094538845422$. Below is the plot of the parabola and the data.



3. If A is a $m \times n$ matrix ($m \geq n$) and it is not full rank (rank $(A) = r < n$), what is the dimension of the set of vectors $\mathbf{x}$ that minimize $\|A\mathbf{x} - \mathbf{b}\|_2$? Why?

   Solution:

   It is the dimension of nullspace: $n - r$. This is because if $\mathbf{x}$ minimizes $\|A\mathbf{x} - \mathbf{b}\|_2$, so does $\mathbf{x} + \alpha\mathbf{z}$, where $A\mathbf{z} = 0$.

4. We will explore solving a least squares problem with the QR factorization. Generate a pseudorandom 10×3 matrix and 10×1 vector in Matlab/Octave as follows:

```
rng(12);
A = rand(10,3);
b = rand(10,1);
rank(A);  % note A has full rank (rank = 3)
```

   (a) Let Matlab find the LS solution with backslash:

```
x = A \ b;
```

   Now find the solution yourself using the QR decomposition of A. Get the QR decomposition as follows:

```
[Q,R] = qr(A);
```

   Write Matlab/Octave code that uses Q and R to find the same solution x that Matlab determined. Include your code and results from the Matlab LS solve and your LS solve.

   (b) Now construct a matrix As which is rank-deficient as follows.

```
R(3,3) = 0;
As = Q*R; % make As a matrix with linearly dependent columns
rank(As); % note As is rank-deficient (rank = 2)
```

What solution does Matlab give in this case? Find this solution yourself with the QR decomposition of As.

Solution:

One possible octave solution is:

```
rand("seed",12);
A = rand(10,3);
b = rand(10,1);
rank(A)

% (a)
x1 = A \ b
[Q,R] = qr(A);
x2 = R \ (Q'*b)
norm(x1-x2) % show that the Matlab solution and the qr solution are very close

% (b)
R(3,3) = 0;
As = Q*R;
rank(As)

xs1 = As \ b

[Q,R] = qr(As)
xs2 = R \ (Q'*b)
```

which solves the matrix A with $x \approx \begin{pmatrix} 0.9648 \\ 0.4001 \\ -0.3835 \end{pmatrix}$ and matrix As with $xs \approx \begin{pmatrix} 0.6993 \\ 0.3219 \\ 0 \end{pmatrix}$.

5. The QR factorization is useful for solving least squares problems when the matrix has full rank. In class, we showed that the SVD can be used to find the minimum norm least squares solution when the matrix does not have full rank. While this works well, it is rather expensive due to the cost of computing the SVD. In this problem we will work out a more efficient solution. The complete orthogonal factorization of A has the form

$$A = U \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T,$$

where U and V are orthogonal matrices and R is invertible and upper triangular.

(a) Devise an algorithm for computing the factorization efficiently. (Hint: this is straightforward to do using QR.)
Solution:
Compute $V, \begin{pmatrix} B \\ \mathbf{z} \end{pmatrix} \leftarrow qr(A^T)$ and $U, \begin{pmatrix} R \\ \mathbf{z} \end{pmatrix} \leftarrow qr(B^T)$.

(b) Show that the factorization above can be used to solve the minimum norm least squares problem $\|A\mathbf{x} - \mathbf{b}\|_2$.
Solution:

$$A^T A\mathbf{x} = A^T \mathbf{b}$$

$$V \begin{pmatrix} R^T & 0 \\ 0 & 0 \end{pmatrix} U^T U \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T \mathbf{x} = V \begin{pmatrix} R^T & 0 \\ 0 & 0 \end{pmatrix} U^T \mathbf{b}$$

$$\begin{pmatrix} R^T & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T \mathbf{x} = \begin{pmatrix} R^T & 0 \\ 0 & 0 \end{pmatrix} U^T \mathbf{b}$$

Let

$$\begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix} = V^T \mathbf{x}$$

$$\begin{pmatrix} \mathbf{c} \\ \mathbf{d} \end{pmatrix} = U^T \mathbf{b}$$

Then the non-zero part, $\mathbf{y}$, of the solution vector is given by

$$R^T R \mathbf{y} = R^T \mathbf{c}$$

Therefore, we obtain the minimum norm solution

$$\mathbf{x} = V \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}.$$

Eigenvalues and Eigenvectors

6. (Heath 4.3) Given an approximate eigenvector $\mathbf{x}$ of A, what is the best estimate (in the least squares sense) for the corresponding eigenvalue?

Solution:

Given $\mathbf{x} \neq \mathbf{0}$ and A, find λ that minimizes

$$\|\mathbf{x}\lambda - A\mathbf{x}\|_2$$

. Form the normal equations:

$$\mathbf{x}^T \mathbf{x} \lambda = \mathbf{x}^T A \mathbf{x}$$

$$\Rightarrow \lambda = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$$

.

7. (Heath 4.1)

(a) Prove that 5 is an eigenvalue of the matrix

$$A = \begin{pmatrix} 6 & 3 & 3 & 1 \\ 0 & 7 & 4 & 5 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Solution:

λ is an eigenvalue if $\det(A - \lambda I) = 0$.

$$\det(A - 5I) = \det \begin{pmatrix} 1 & 3 & 3 & 1 \\ 0 & 2 & 4 & 5 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 3 \end{pmatrix} = 0,$$

since $A - 5I$ is a triangular matrix with a 0 diagonal element.

(b) Exhibit an eigenvector of A corresponding to the eigenvalue 5. **Solution:**

We need to find a null vector $\mathbf{x} = (x, y, z, w)$ of $A - 5I$. This can be done by working backwards. Choose $w = 0$. Then the last two elements of $(A - 5I)\mathbf{x}$ are 0. Choose $y = 2$, and $z = -1$. Then the second element of $(A - 5I)\mathbf{x}$ is also 0. Finally, we must have $x = -3$ so that the first element of $(A - 5I)\mathbf{x}$ is 0. Check:

$$\begin{pmatrix} 6 & 3 & 3 & 1 \\ 0 & 7 & 4 & 5 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \\ -1 \\ 0 \end{pmatrix} = 5 \begin{pmatrix} -3 \\ 2 \\ -1 \\ 0 \end{pmatrix} \checkmark$$

8. Let $A\mathbf{u} = \lambda\mathbf{u}$ for some eigenvalue λ and eigenvector $\mathbf{u}$. Find an eigenvalue/eigenvector pair for each of the following:

   (a) A^{-1}
   $A\mathbf{u} = \lambda\mathbf{u} \implies \lambda^{-1}\mathbf{u} = A^{-1}\mathbf{u}$. Eigenvector: $\mathbf{u}$; eigenvalue: λ^{-1}.

   (b) cA
   $(cA)\mathbf{u} = (c\lambda)\mathbf{u}$. Eigenvector: $\mathbf{u}$; eigenvalue: $c\lambda$.

   (c) $A + cI$
   $(A + cI)\mathbf{u} = (\lambda + c)\mathbf{u}$. Eigenvector: $\mathbf{u}$; eigenvalue: $\lambda + c$.

   (d) A^2
   $A^2\mathbf{u} = \lambda A\mathbf{u} = \lambda^2\mathbf{u}$. Eigenvector: $\mathbf{u}$; eigenvalue: λ^2.

   (e) $(A + 2I)(A - I)^2$
   Eigenvector: $\mathbf{u}$; eigenvalue: $(\lambda + 2)(\lambda - 1)^2$.

9. Let $B = A^T A$, $C = AA^T$, and $B\mathbf{u} = \lambda\mathbf{u}$.

   (a) Find an eigenvalue/eigenvector pair for C. You may assume $\lambda \neq 0$.
   Solution:

$$\begin{aligned} A^T A\mathbf{u} &= \lambda\mathbf{u} \\ AA^T A\mathbf{u} &= \lambda A\mathbf{u} \\ AA^T\mathbf{v} &= \lambda\mathbf{v} \qquad \mathbf{v} = A\mathbf{u} \end{aligned}$$

   Note that this only works provided $\mathbf{v} \neq 0$. Since $A^T\mathbf{v} = \lambda\mathbf{u}$, $\lambda \neq 0$ ensures this.

   (b) What happens if instead $\lambda = 0$?
   Solution:
   In this case, the given eigenvalue/eigenvector pair for B might not have a corresponding pair for C.

10. Let Q be an orthogonal matrix and $Q\mathbf{u} = \lambda\mathbf{u}$. Show that $|\lambda| = 1$. Note that λ will in general be complex. (Hint: it is easier to assume that Q unitary instead. It is actually easier to treat Q as possibly complex-valued.) **Solution:**

$$\begin{aligned} Q\mathbf{u} &= \lambda\mathbf{u} \\ (Q\mathbf{u})^H(Q\mathbf{u}) &= (\lambda\mathbf{u})^H(\lambda\mathbf{u}) \\ \mathbf{u}^H Q^H Q\mathbf{u} &= \overline{\lambda}\lambda\mathbf{u}^H\mathbf{u} \\ \mathbf{u}^H\mathbf{u} &= |\lambda|^2\mathbf{u}^H\mathbf{u} \\ 1 &= |\lambda|^2 \\ 1 &= |\lambda| \end{aligned}$$