AMATH 352: Problem Set 8 Solutions

Dave Moore, dmmfix@uw.edu

March 4, 2017

Due: Friday March 10, 2017

Least Squares:

1. Find the least squares solutions to the following linear systems by finding and solving the normal equations by hand (use Gaussian elimination). Show your work. You can check your answers by verifying that the residual is orthogonal to the columns of $\boldsymbol{A}$, i.e. that $\boldsymbol{A}^T\boldsymbol{r} = \boldsymbol{0}$.

   (a) $\boldsymbol{A} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

   (b) $\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ 4 & -2 \\ 1 & 5 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

Solution:

(a) We want $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ so

$$\begin{bmatrix} 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} x = \begin{bmatrix} 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
$$11x = 4$$
$$x = \frac{4}{11}$$

Verifying:

$$\boldsymbol{r} = \frac{4}{11} \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{7}{11} \\ \frac{1}{11} \\ -\frac{4}{11} \end{bmatrix}$$
$$\boldsymbol{A}^T\boldsymbol{r} = \begin{bmatrix} 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} -\frac{7}{11} \\ \frac{1}{11} \\ -\frac{4}{11} \end{bmatrix} = \frac{-7+3+4}{11} = 0$$

(b)

$$\begin{bmatrix} 2 & 4 & 1 \\ 3 & -2 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & -2 \\ 1 & 5 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 21 & 3 \\ 3 & 38 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 13 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{7} & 1 \end{bmatrix} \begin{bmatrix} 21 & 3 \\ 3 & 38 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{7} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 13 \end{bmatrix}$$

$$\begin{bmatrix} 21 & 3 \\ 0 & 37\frac{4}{7} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 12\frac{6}{7} \end{bmatrix}$$

Solving the upper-triangular system by back-substitution

$$37\frac{4}{7}\boldsymbol{x}_2 = 12\frac{6}{7} \implies \boldsymbol{x}_2 = 0.3422$$

$$21\boldsymbol{x}_1 + 3(0.3422) = 1 \implies \boldsymbol{x}_1 = -0.001266$$

Verifying:

$$\boldsymbol{r} = \begin{bmatrix} 2 & 3 \\ 4 & -2 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} -0.001266 \\ 0.3422 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.9759 \\ 0.3105 \\ 0.7098 \end{bmatrix}$$

$$\boldsymbol{A}^T\boldsymbol{r} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & -2 & 5 \end{bmatrix} \begin{bmatrix} -0.9759 \\ 0.3105 \\ 0.7098 \end{bmatrix} = \mathbf{0}$$

2. **Linear regression:** For the following data $\begin{array}{c|cccc} t_i & 1 & 2 & 3 & 5 \\ \hline y_i & 1 & 0 & -2 & -3 \end{array}$, find the line $y = \alpha + \beta t$ that best fits the data in the least squares sense. You may use Matlab to help you solve this problem, but in your writeup you must derive the linear system $\boldsymbol{Ac} = \boldsymbol{b}$ that needs to be approximately solved. You do not need to include the code used to solve this problem in the script you submit to Scorelator.

Solution:

We are finding a best fit curve in $\mathbb{P}_1$, we'll use the standard basis $\{1, t\}$ for the space. So we need to find the least-squares solution to the following linear system:

$$\min_{\boldsymbol{x}} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 = \min_{\boldsymbol{x}} \left\| \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix} \boldsymbol{x} - \begin{bmatrix} 1 \\ 0 \\ -2 \\ -3 \end{bmatrix} \right\|_2^2$$

The columns of $\boldsymbol{A}$ are simply the basis functions evaluated at t_i. Solving with backslash in Matlab, we obtain

$$\boldsymbol{x} = \begin{bmatrix} 0.2286 \\ -0.6286 \end{bmatrix}$$

3. One can also use the least squares methodology to approximate data using basis functions which are more general than polynomials. For example, one can also use trigonometric functions. Given the data $\begin{array}{c|ccc} t_i & 0 & 0.5 & 1 \\ \hline y_i & 1 & 0.5 & 0.25 \end{array}$, find the trigonometric function of the form $g(t) = \alpha\cos(\pi t) + \beta\sin(\pi t)$ that best approximates the data in the least squares sense. To do this find and solve the normal equations by hand.

Solution:

As in Problem 2, we want to find a least-squares solution for a system in which the columns of our matrix are simply the basis functions evaluated at t_i, which yields the following system:

$$\boldsymbol{A} = \begin{bmatrix} \cos(0) & \sin(0) \\ \cos(\pi/2) & \sin(\pi/2) \\ \cos(\pi) & \sin(\pi) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} 1 \\ 0.5 \\ 0.25 \end{bmatrix}$$

The normal equations for this system are

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

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0.5 \\ 0.25 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.5 \end{bmatrix}$$

Since the matrix on the left is diagonal, we can read off the solution as

$$\alpha = 0.375$$
$$\beta = 0.5$$

4. Consider the least squares problem

$$\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2.$$

(a) Show that if $\hat{\boldsymbol{x}}$ minimizes $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$, then so does $\hat{\boldsymbol{x}} + \boldsymbol{w}$ for any $\boldsymbol{w} \in N(\boldsymbol{A})$. Here we have dropped the assumption from lecture that $N(\boldsymbol{A}) = \{\boldsymbol{0}\}$.

(b) Show that if $\boldsymbol{u}$ and $\boldsymbol{v}$ are both least-squares solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, then $\boldsymbol{u} - \boldsymbol{v} \in N(\boldsymbol{A})$.

(c) Suppose now that $\boldsymbol{A}$ satisfies $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$. In this case, you can write down a closed form solution for $\hat{\boldsymbol{x}}$. Derive this solution. Note that $\boldsymbol{A}$ is not necessarily orthogonal since it may not be square, e.g. $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ satisfies $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$, but it is not orthogonal.

Solution:

(a) Since $\boldsymbol{w} \in N(\boldsymbol{A})$, we can see that

$$\|\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{w}) - \boldsymbol{b}\|_2^2 = \|(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{w}) - \boldsymbol{b}\|_2^2 = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

and so the norms of the residuals for $\hat{\boldsymbol{x}} + \boldsymbol{w}$ and $\boldsymbol{x}$ are equivalent.

(b) Since a least-squares solution is the unique point in range($\boldsymbol{A}$) that minimizes $\|\boldsymbol{r}\|_2^2$, if $\boldsymbol{u}$ and $\boldsymbol{v}$ minimize $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$, we must have

$$\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b} = \boldsymbol{r}$$
$$\boldsymbol{A}\boldsymbol{v} - \boldsymbol{b} = \boldsymbol{r}$$

By subtracting these we find

$$\boldsymbol{A}\boldsymbol{u} - \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$$
$$\boldsymbol{A}(\boldsymbol{u} - \boldsymbol{v}) = \boldsymbol{0}$$

Or in otherwords, $\boldsymbol{u} - \boldsymbol{v} \in N(\boldsymbol{A})$.

(c) In this special case, we can use the normal equations to derive a solution xfor x:

$$\begin{aligned} \boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{x}} &= \boldsymbol{A}^T\boldsymbol{b} \\ \boldsymbol{I}\hat{\boldsymbol{x}} &= \boldsymbol{A}^T\boldsymbol{b} \\ \hat{\boldsymbol{x}} &= \boldsymbol{A}^T\boldsymbol{b} \end{aligned}$$

5. Suppose $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $m \geq n$, $\text{rank}(\boldsymbol{A}) = n$, and we have access to the reduced singular value decomposition (SVD) of $\boldsymbol{A}$:

$$\boldsymbol{A} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T.$$

Here $\hat{\boldsymbol{U}} \in \mathbb{R}^{m\times n}$ has orthonormal columns (so $\hat{\boldsymbol{U}}^T\hat{\boldsymbol{U}} = \boldsymbol{I} \in \mathbb{R}^{n\times n}$, but $\hat{\boldsymbol{U}}$ is not orthogonal unless it is square), $\hat{\boldsymbol{\Sigma}} \in \mathbb{R}^{n\times n}$ is a diagonal matrix with positive entries on the diagonal, and $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix.

Derive a method of solving the least squares problem $\min_x \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$ using the SVD of $\boldsymbol{A}$. You should end up with a linear equation to solve which involves at most one instance of $\hat{\boldsymbol{U}}$, $\hat{\boldsymbol{\Sigma}}$, and $\boldsymbol{V}$. Hint: plug the SVD of $\boldsymbol{A}$ into the normal equations.

Solution:

Let's replace $\boldsymbol{A}$ with its SVD in the normal equations. Note that we can apply both of our existing identities $\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{I}$ and $\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{I}$ to reduce both sides. We also know that $\hat{\boldsymbol{\Sigma}}$ is non-singular and therefore invertible since it is a diagonal matrix with non-zero entries, which allows us to remove it from both sides as needed. We also observe that $\hat{\boldsymbol{\Sigma}}^T = \hat{\boldsymbol{\Sigma}}$.

$$\begin{aligned} \boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{x}} &= \boldsymbol{A}^T\boldsymbol{b} \\ (\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T)^T\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T\boldsymbol{x} &= (\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T)^T\boldsymbol{b} \\ \boldsymbol{V}\hat{\boldsymbol{\Sigma}}^T\hat{\boldsymbol{U}}^T\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T\boldsymbol{x} &= \boldsymbol{V}\hat{\boldsymbol{\Sigma}}^T\hat{\boldsymbol{U}}^T\boldsymbol{b} \\ \boldsymbol{V}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{U}}^T\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T\boldsymbol{x} &= \boldsymbol{V}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{U}}^T\boldsymbol{b} \\ \boldsymbol{V}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T\boldsymbol{x} &= \boldsymbol{V}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{U}}^T\boldsymbol{b} \\ \hat{\boldsymbol{\Sigma}}\boldsymbol{V}^T\boldsymbol{x} &= \hat{\boldsymbol{U}}^T\boldsymbol{b} \\ \boldsymbol{V}^T\boldsymbol{x} &= \hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{U}}^T\boldsymbol{b} \\ \boldsymbol{x} &= \boldsymbol{V}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{U}}^T\boldsymbol{b} \end{aligned}$$