Linear Algebra Homework 6

William Jardee

May 2020

1. (a) Find the lease squares solution

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix}$$

Least squares solution is $A^T A\mathbf{x} = A^T\mathbf{b} \rightarrow \mathbf{x} = (A^T A)^{-1} A^T\mathbf{b}$

$A^T A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$

$(A^T A)^{-1} = \frac{1}{42-36} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}$

$(A^T A)A^T = \frac{1}{6} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -3 & 0 & 3 \\ 8 & 2 & -4 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix} =$
$\frac{1}{6} \begin{bmatrix} -15 \\ 50 \end{bmatrix} \approx \begin{bmatrix} -2.5 \\ 8.33 \end{bmatrix}$

So the least squares line is: $y = -2.5x + 8.33$.

(b) The sum of squares of error, or SSE, is calculated with $\|A\mathbf{x} - \mathbf{b}\|^2$.

$A\mathbf{x} \approx \begin{bmatrix} 5.833 \\ 3.33 \\ 0.833 \end{bmatrix}$

$SSE(\mathbf{x}) = \left\| \begin{bmatrix} 5.83 \\ 3.33 \\ 0.833 \end{bmatrix} - \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} -0.167 \\ 0.33 \\ -0.167 \end{bmatrix} \right\|^2 \approx 0.167$

(c) Since A has linearly independent columns, $(A^T A)^{-1}$ exists and is unique. Consequently the least squares solution is unique.

2. (a) Find the least squares solution

$$y = a + bx + cx^2$$

$$\begin{bmatrix}1 & 1 & 1\\1 & 2 & 4\\1 & 3 & 9\\1 & 4 & 16\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=\begin{bmatrix}6\\0\\0\\2\end{bmatrix}$$

Again, $\mathbf{x}_{l.s.} = (A^TA)^{-1}A^T\mathbf{b}$

$A^TA = \begin{bmatrix}4 & 10 & 30\\10 & 30 & 100\\30 & 100 & 354\end{bmatrix}$

The inverse was too extensive to do by hand reliably, so with assistance from Matlab:

$(A^TA)^{-1} \approx \begin{bmatrix}7.75 & -6.75 & 1.25\\-6.75 & 6.45 & -1.25\\1.25 & -1.25 & 0.25\end{bmatrix}$

$(A^TA)^{-1}A^T \approx \begin{bmatrix}2.25 & -0.75 & -1.25 & 0.75\\-1.55 & 1.15 & 1.35 & -0.95\\0.25 & -0.25 & -0.25 & 0.25\end{bmatrix}$

$(A^TA)^{-1}A^T\mathbf{b} \approx \begin{bmatrix}15\\-11.2\\2\end{bmatrix}$

So the least squares line approximation is: $y = 15 - 11.2x + 2x^2$

(b) $SSE(\mathbf{x}) = \|A\mathbf{x}-\mathbf{b}\|^2 \approx \left\|\begin{bmatrix}5.8\\0.6\\-0.6\\2.2\end{bmatrix}-\begin{bmatrix}6\\0\\0\\2\end{bmatrix}\right\|^2 \approx 0.8$

3.

$$A=\begin{bmatrix}0 & 1 & 1 & 0\\1 & -1 & 1 & -1\\1 & 0 & 1 & 0\\1 & 1 & 1 & 1\end{bmatrix}\quad b=\begin{bmatrix}5\\3\\-1\\1\end{bmatrix}$$

(a) If a matrix doesn't have linearly independent columns, the solutions either don't exist or are not unique. The least squares solution is given by $A^TA\mathbf{x} = A^T\mathbf{b}$.

$$A^TA=\begin{bmatrix}3 & 0 & 3 & 0\\0 & 3 & 1 & 2\\3 & 1 & 4 & 0\\0 & 2 & 0 & 2\end{bmatrix}$$

The 3rd row can be written as a linear combination of the other three rows:

$$\begin{bmatrix}3\\0\\3\\0\end{bmatrix}+\begin{bmatrix}0\\3\\1\\2\end{bmatrix}-\begin{bmatrix}0\\2\\0\\2\end{bmatrix}=\begin{bmatrix}3\\1\\4\\0\end{bmatrix}$$

This means that if we solve $(A^TA)\mathbf{x} = A^T\mathbf{b}$, there is not one unique solution.

(b) To show there are an infinite number of least squares solutions, we have to show that there are neither 0 or 1 solution. We have already ruled out 1 solution. To rule out 0 solutions, we can look at what it means to be a least squares solution. The least squares solution solves for the projection of $\mathbf{b}$ onto A. If we take the a basis of A, say U, then $Proj_A(\mathbf{b}) = \mathbf{b}\cdot\mathbf{u}_1+\mathbf{b}\cdot\mathbf{u}_2+\cdots+\mathbf{b}\cdot\mathbf{u}_n = \mathbf{x}_{l.s.}$. So $\mathbf{x}_{l.s.}$ always has at least one solution, since the solution cannot be undefined. We get multiple solutions when the SSE of multiple of those projections are the same.

4. (a) We need to show $P^T = P$
$P^T = (A(A^TA)^{-1}A^T)^T = (A^T)^T((A^TA)^{-1})^TA^T = A((A^TA)^T)^{-1}A^T =$
$A(A^TA)^{-1}A^T = P$
So P is symmetric

(b) We need to show $P^2 = P$
$P^2 = PP = (A(A^TA)^{-1}A^T)(A(A^TA)^{-1}A^T) = A(A^TA)^{-1}(A^TA)(A^TA)^{-1}A^T =$
$A((A^TA)^{-1})A^T = P$
So P is idempotent

5. Taking the assumption that A is square and has linearly independent columns, we can deduce that the A is invertible. Then the projection matrix P is:

$$P = A(A^TA)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I\cdot I = I$$

.

6. Prove that if a matrix A has independent columns (and possibly rectangular!), then A^TA is invertible.

Proof:

If we show that 0 is not an eigenvector of A^TA, then it is invertible. If A has linear independent columns, then there is no null-space, $A\mathbf{x} \neq \mathbf{0}$ $\forall$ $\mathbf{x} \in \mathbb{R}^n$ and consequently $\|A\mathbf{x}\|^2 \neq 0$. Let us $\mathbf{v}$ to be an eigenvector of A^TA.

$$0 \neq \|A\mathbf{v}\|^2 = A\mathbf{v}\cdot A\mathbf{v} = \mathbf{v}^TA^TA\mathbf{v} = \mathbf{v}^T\lambda\mathbf{v} = \lambda(\mathbf{v}\cdot\mathbf{v}) = \lambda\|\mathbf{x}\|^2$$

Since the definition of eigenvector doesn't allow $\mathbf{v}$ to be the zero vector, $\|\mathbf{v}\|^2 \neq 0$, so $\lambda \neq 0$. Since $\lambda \neq 0$, then A^TA is invertible.

"Quack" ■

7. Prove that if a square matrix B is diagonalizable with r non-zero eigenvalues, then B has rank r.

Proof:

If A is a $n \times n$ diagonalizable matrix, then it has all linearly independent eigenvectors. The eigenvectors that correlate to the eigenvalues of zero form a basis for the null-space of A. If there are r non-zero eigenvalues, then there are n-r zero eigenvalues, and consequently n-r vectors that make up the basis for our null-space. The $nullity(A) = n - r$. We know through the Rank Theorem that $n = rank + nullity(A)$. If we substitute in that $nullity(A) = n - r$: $n = rank + n - r \rightarrow r = rank$.

"Quack" ■

8. Let $A = U\Sigma V^T$ be the SVD of a (possibly rectangular) matrix A. Prove that the columns of U are orthogonal.

Proof:

If we let $A = U\Sigma V^T$, by the definition of a SVD we know that the columns of U are either $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$, or found using gram-schmitt to finish the basis. Using gram-schmitt we are guaranteed to have orthogonal vectors, so the vectors past the r^{th} vector will be be orthogonal to every vector in the matrix. Looking at the first r vectors:

$$\mathbf{u}_i \cdot \mathbf{u}_j = (\frac{1}{\sigma_i} A\mathbf{v}_i) \cdot (\frac{1}{\sigma_j} A\mathbf{v}_j) = \frac{1}{\sigma_i \sigma_j} \mathbf{v}_i^T A^T A \mathbf{v}_j = \frac{1}{\sigma_i \sigma_j} \mathbf{v}_i^T \lambda_j \mathbf{v}_j = \frac{\lambda_j}{\sigma_i \sigma_j} \mathbf{v}_i \cdot \mathbf{v}_j$$

If $i \neq j$ then $\mathbf{v}_i \cdot \mathbf{v}_j = 0 = \mathbf{u}_i \cdot \mathbf{u}_j$. We see that any two column vectors of U are orthogonal.

"Quack" ■

9. Let A be a symmetric matrix. Show that the singular values of A are the absolute values of the eigenvalues of A.

Proof:

For a matrix A, if A is symmetric, then it is orthogonally diagonalizable. $A = S\Lambda S^T$, $A^T A = A^2 = (S\Lambda S^T)(S\Lambda S^T) = S\Lambda^2 S^T$.
This is the orthogonal diagonalization of $A^T A$, so the eigenvalues of $A^T A$ are the squares of the eigenvalues of A. The singular values are defined to be the principle root of the eigenvalues of $A^T A$, so the singular values are $\sigma_i = \left| \sqrt{\lambda_i^2} \right| = |\lambda|_i$. So the singular values of A are the absolute values of the eigenvalues of A.

"Quack" ■

10. Show that if $A = U\Sigma V^T$ is an SVD of A, then the left singular vectors are eigenvectors of AA^T.

Proof:

$$AA^T \mathbf{u}_i = AA^T \frac{1}{\sigma_i} A\mathbf{v}_i = \frac{1}{\sigma_i} AA^T A\mathbf{v}_i = \frac{1}{\sigma_i} A\lambda_i \mathbf{v}_i = \frac{\lambda_i}{\sigma_i} A\mathbf{v}_i = \lambda_i \mathbf{u}_i$$

So, for when $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$, which is for all the left singular vectors, $\mathbf{u}_i$ is an eigenvector with an eigenvalue of λ_i, which is the "i^{th}" singular value squared.

"Quack" $\blacksquare$