Questions

1. In lecture 5, we showed that the solution of the least squares regression problem used the pseudoinverse:

$$\mathbf{A}^{+} = \mathbf{A}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{A}^{\mathbf{T}}.$$

   Using the singular value decomposition of $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathbf{T}}$, show that

$$\mathbf{A}^{+} = \mathbf{V}\mathbf{\Sigma}^{-\mathbf{1}}\mathbf{U}^{\mathbf{T}}$$

   To do this correctly, you will need to review how transposes and inverses work, and also keep dimensions in mind.

2. Show that

$$\mathbf{A}\mathbf{A}^{+} = \mathbf{U}\mathbf{U}^{\mathbf{T}}$$

   Note that $\mathbf{U}\mathbf{U}^{\mathbf{T}}$ is an $m \times m$ matrix, when $\mathbf{U}$ is $m \times n$.

   What can you say about the elements of matrix? (In the lecture video, I mistakenly claimed that it was diagonal, and then had to correct that slide afterwards.)

Answers

1.

$$\begin{aligned}
\mathbf{A}^+ &= (\mathbf{A^T A})^{-\mathbf{1}}\mathbf{A^T} \\
&= ((\mathbf{U\Sigma V^T})^{\mathbf{T}}(\mathbf{U\Sigma V^T}))^{-\mathbf{1}}(\mathbf{U\Sigma V^T})^{\mathbf{T}} \\
&= ((\mathbf{V\Sigma U^T U\Sigma V^T})^{-\mathbf{1}}\mathbf{V\Sigma U} \\
&= ((\mathbf{V\Sigma^2 V^T})^{-\mathbf{1}}\mathbf{V\Sigma U}, \text{ since } \mathbf{U^T U} = \mathbf{I} \\
&= \mathbf{V\Sigma^{-2} V^T V\Sigma U}, \text{ since } \mathbf{V^{-1}} = \mathbf{V^T} \\
&= \mathbf{V\Sigma^{-1} U}
\end{aligned}$$

2. From the previous question, we have

$$\begin{aligned}
\mathbf{AA}^+ &= (\mathbf{U\Sigma V^T})(\mathbf{V\Sigma^{-1} U^T}) \\
&= \mathbf{U\Sigma\Sigma^{-1} U^T} \\
&= \mathbf{UU^T}
\end{aligned}$$

Note that when $m > n$, the matrix $\mathbf{UU^T}$ is not the identity. The reason is that $\mathbf{U}$ only has n independent columns.

To understand what the element of this matrix are, note that $\mathbf{UU^T x} = \mathbf{x}$ when $\mathbf{x}$ is a column of $\mathbf{U}$ so it behaves like the identity matrix. However, $\mathbf{UU^T x} = \mathbf{0}$ when $\mathbf{x}$ is orthogonal to all the columns of $\mathbf{U}$ (or rows of $\mathbf{U^T}$), so clearly it is not the identity matrix!

