DSC 40A - (Mini) Homework 04

Due: Friday, February 07, 2020

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Unless otherwise noted by the problem's instructions, show your work or provide some justification for your answer. Homeworks are due via Gradescope on Friday afternoon at 5:00 p.m.

Note: This homework is intended to be shorter and easier due to the midterm. Because of this, it will be worth fewer points than a typical homework.

Problem 1.

Let X be an $n \times d$ matrix, let A be a $d \times r$ matrix, and let B be an $r \times n$ matrix. Let $\vec{x}$ be a vector in $\mathbb{R}^n$ (that is, an $n \times 1$ column vector), let $\vec{y}$ be a vector in $\vec{R}^d$ (that is, a $d \times 1$ column vector). For each of the following, state whether the result is a scalar, a vector, or a matrix. If it is a vector or a matrix, state its shape (number of rows and columns).

For the purposes of this question, a matrix with one column is considered a column vector, and a matrix with one row is considered a row vector. If the result of an expression is 1×1, it is a scalar. You do not need to show your work.

a) $\vec{x} \cdot \vec{x}$

Solution: Scalar.

b) XA

Solution: $n \times r$ matrix.

c) $XX^\intercal$

Solution: $n \times n$

d) $X^\intercal X$

Solution: $d \times d$

e) $(XA)^\intercal \vec{x}$

Solution: $r \times 1$ vector.

f) $\vec{y}^\intercal \vec{y}(XX^\intercal)^{-1}$

Solution: $n \times n$ matrix.

g) $(\vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y}) + x^\intercal B^\intercal A^\intercal X^\intercal XAB\vec{x}$

Solution: Scalar.

h) $B^\intercal A^\intercal X^\intercal X A B$

> **Solution:** $n \times n$ matrix.

Problem 2.

This problem will check that we're all on the same page when it comes to the notation used in lecture.

The table below shows data we've collected on the salaries of three data scientists (don't worry, the Salary column is in thousands of dollars).

Person	GPA	Experience	Salary
1	3.3	4	95
2	3.9	10	120
3	3.2	3	80

Suppose we have decided on a prediction rule:

$$H(\vec{x}) = 50 + 10 \times x_1 + 2 \times x_2,$$

where the first component of $\vec{x}$, x_1, represents GPA and the second component, x_2, represents Experience.

a) Write down the parameter vector, $\vec{w}$. Assume that it includes w_0. Your answer should be a vector with three elements.

> **Solution:**
> $$\vec{w} = \begin{pmatrix} 50 \\ 10 \\ 2 \end{pmatrix}$$

b) Write down the data vectors $\vec{x}^{(1)}$, $\vec{x}^{(2)}$, and $\vec{x}^{(3)}$ for the first, second, and third person in data set, respectively.

> **Solution:**
> $$\vec{x}^{(1)} = \begin{pmatrix} 3.3 \\ 4 \end{pmatrix} \qquad \vec{x}^{(2)} = \begin{pmatrix} 3.9 \\ 10 \end{pmatrix} \qquad \vec{x}^{(3)} = \begin{pmatrix} 3.2 \\ 3 \end{pmatrix}$$

c) Compute the predicted salaries $H(\vec{x}^{(1)})$, $H(\vec{x}^{(2)})$, $H(\vec{x}^{(3)})$ for each of the three people in the data set.

Solution:

$$
\begin{aligned}
H(\vec{x}^{(1)}) &= \vec{w} \cdot \text{Augmented}(\vec{x}^{(1)}) \\
&= \begin{pmatrix} 50 \\ 10 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3.3 \\ 4 \end{pmatrix} \\
&= 50 + 33 + 8 \\
&= 91
\end{aligned}
$$

$$
\begin{aligned}
H(\vec{x}^{(2)}) &= \vec{w} \cdot \text{Augmented}(\vec{x}^{(2)}) \\
&= \begin{pmatrix} 50 \\ 10 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3.9 \\ 10 \end{pmatrix} \\
&= 50 + 39 + 20 \\
&= 109
\end{aligned}
$$

$$
\begin{aligned}
H(\vec{x}^{(3)}) &= \vec{w} \cdot \text{Augmented}(\vec{x}^{(3)}) \\
&= \begin{pmatrix} 5010 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3.2 \\ 3 \end{pmatrix} \\
&= 50 + 32 + 6 \\
&= 88
\end{aligned}
$$

d) Compute the mean squared error of this prediction rule (with this particular choice of parameters).

Solution: The squared error of each prediction is:

$$
\begin{aligned}
(H(\vec{x}^{(1)}) - y_1)^2 &= (91 - 95)^2 = 16 \\
(H(\vec{x}^{(2)}) - y_2)^2 &= (109 - 120)^2 = 121 \\
(H(\vec{x}^{(3)}) - y_3)^2 &= (88 - 80)^2 = 64
\end{aligned}
$$

That makes the mean squared error:

$$(16 + 121 + 64)/3 = 201 = 67.$$

e) Write down the *design matrix*, X.

Solution:

$$X = \begin{pmatrix} 1 & 3.3 & 4 \\ 1 & 3.9 & 10 \\ 1 & 3.2 & 3 \end{pmatrix}$$

f) Check that the entries of $X\vec{w}$ are the predicted salaries you found above.

Solution:

$$
\begin{aligned}
X\vec{w} &= \begin{pmatrix} 1 & 3.3 & 4 \\ 1 & 3.9 & 10 \\ 1 & 3.2 & 3 \end{pmatrix} \begin{pmatrix} 50 \\ 10 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} 50 + 33 + 8 \\ 50 + 39 + 20 \\ 50 + 32 + 6 \end{pmatrix} \\
&= \begin{pmatrix} 91 \\ 109 \\ 88 \end{pmatrix}
\end{aligned}
$$

g) Calculate the norm of the vector $X\vec{w} - \vec{y}$, where $\vec{y} = (95, 120, 80)^\intercal$ is the vector of observations.

Solution:

$$
\begin{aligned}
\|X\vec{w} - \vec{y}\| &= \left\| \begin{pmatrix} 91 \\ 109 \\ 88 \end{pmatrix} - \begin{pmatrix} 95 \\ 120 \\ 80 \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} 4 \\ 11 \\ 8 \end{pmatrix} \right\| \\
&= \sqrt{4^2 + 11^2 + 8^2} \\
&= \sqrt{16 + 121 + 64} \\
&= \sqrt{201}
\end{aligned}
$$

h) Check that 1/3 of the *squared norm* of $X\vec{w} - \vec{y}$ is the mean squared error you found above.

Solution:

$$
\frac{1}{2} \cdot \sqrt{201}^2 = \frac{201}{3} = 67
$$