DSC 140A - Math Check Worksheet

One of the things that can make machine learning theory challenging to learn is that it builds upon a lot of different math classes that you might have taken a long time ago – classes like multivariate calculus, linear algebra, and probability. In DSC 140A, we'll spend some time in lecture reviewing the background math, but we won't have time to cover everything in depth.

This worksheet is meant to be a way to give yourself a quick diagnostic check on your math background. **It's not graded, and it won't be turned in**. Instead, the idea is that you can use it to see what you remember and what you've forgotten, and to use the results to see what you might want to brush up on.

This worksheet is paired with a set of slides that cover the background math in more detail. These slides list a bunch of important "Facts" that you probably saw in a previous class, but may have forgotten. (Almost) every question below will use one or more of these facts. We'll also be using these facts regularly in the rest of the course.

You can find the slides at:

`http://dsc140a.com/materials/default/supplementary/math_review/slides.pdf`

Instructions:

Here's how we recommend using this worksheet:

1. Try each of the problems below without looking at the solutions.
2. Check the solutions to see if you got the right answer. You can find the solutions at

   `http://dsc140a.com/materials/default/supplementary/math_review/solution.pdf`.
3. If you got the wrong answer or just feel like you need more practice, look up the relevant part in the slides. Each problem below uses one or more important "Facts" that you probably learned somewhere in a previous class. The associated slides list these facts and explain them. If you're still unsure after taking a look at the slides, feel free to come to office hours (preferred) or ask on Campuswire!
4. Do a quick scan of the slides to see all of the "Facts", since some of them might not have been used in the problems below.

1 Summation Notation

Problem 1.

True or false: $\sum_{i=1}^n 6(x_i + 10) = (6\sum_{i=1}^n x_i) + 60n$

Solution: True. See Fact 2 and Fact 4 in the math review slides.

To see how you can prove these properties, see Fact 3.

Problem 2.

How should we interpret $\displaystyle\sum_{i=1}^n x_i + y_i$?

● $\sum_{i=1}^n (x_i + y_i)$

○ $(\sum_{i=1}^n x_i) + y_i$

Solution: It *has* to mean $\sum_{i=1}^{n}(x_i + y_i)$, because $(\sum_{i=1}^{n} x_i) + y_i$ does not make notational sense! Just like in programming, i is "unbound" in the second choice, so y_i is not defined!

2 Vectors

Problem 3.

Compute $(1, 4, 3)^T + (2, 0, 1)^T$.

Solution: $(1, 4, 3)^T + (2, 0, 1)^T = (1 + 2, 4 + 0, 3 + 1)^T = (3, 4, 4)^T$. From Fact 6.

Problem 4.

Compute $4(1, 4, 3)^T$.

Solution: $4(1, 4, 3)^T = (4 \cdot 1, 4 \cdot 4, 4 \cdot 3)^T = (4, 16, 12)^T$. From Fact 7.

Problem 5.

Compute $(1, 4, 3)^T \cdot (2, 0, 1)^T$. Here, $\cdot$ denotes the dot product.

Solution: $(1, 4, 3)^T \cdot (2, 0, 1)^T = 1 \cdot 2 + 4 \cdot 0 + 3 \cdot 1 = 5$. From Fact 8.

Problem 6.

Two vectors $\vec{u}$ and $\vec{v}$ are orthogonal to one another (the angle between them is $90°$). What is $\vec{u} \cdot \vec{v}$?

Solution: Zero.

$\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos(90°)$. Since $\cos(90°) = 0$, the dot product is zero. From Fact 9.

Problem 7.

$\vec{u} = (1, 2, 3)^T$. What is the length of $\vec{u}$?

Solution: $\|\vec{u}\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$. From Fact 5.

Problem 8.

Suppose $\vec{v} = (3, 3)^T$.

1. Find a unit vector $\vec{u}^{(1)}$ such that $\vec{u}^{(1)} \cdot \vec{v} = 0$.
2. Find a unit vector $\vec{u}^{(2)}$ such that $|\vec{u}^{(2)} \cdot \vec{v}|$ is maximized.

Solution:

1. $\vec{u}^{(1)} = (1, -1)^T/\sqrt{2}$.
2. $\vec{u}^{(2)} = (1, 1)^T/\sqrt{2}$.

These come from Fact 9. That fact tells us that $\vec{v} \cdot \vec{u}^{(1)} = \|v\| \|\vec{u}^{(1)}\| \cos(\theta)$. $\vec{u}^{(1)}$ is a unit vector, so $\|\vec{u}^{(1)}\| = 1$. Likewise, $\|\vec{v}\|$ is fixed – we can't change it. So to make the dot product zero, we need to make $\cos(\theta) = 0$, which happens when $\theta = 90°$, meaning $\vec{u}^{(1)}$ is orthogonal to $\vec{v}$. You can check that $(1, -1)^T / \sqrt{2}$ is a unit vector, and it's orthogonal to $\vec{v}$.

Likewise, to maximize $|\vec{v} \cdot \vec{u}^{(2)}|$, we need to make $\cos(\theta) = 1$, which happens when $\theta = 0°$. In other words, when $\vec{u}^{(2)}$ is parallel to $\vec{v}$.

Problem 9.

Which of these is another expression for the norm of $\vec{u}$ (that is, $\|\vec{u}\|$)?

- ○ $\vec{u} \cdot \vec{u}$
- ○ $\sqrt{\vec{u}^2}$
- ● $\sqrt{\vec{u} \cdot \vec{u}}$
- ○ $\vec{u}^2$

Solution: This also comes from Fact 5 and/or Fact 9.

Using the coordinate definition of the dot product, we see that the norm of $\vec{u}$ is $\sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_d^2}$, which we recognize as the Euclidean norm.

Another way to see this is to use the geometric definition of the dot product: $\vec{u} \cdot \vec{u} = \|\vec{u}\| \|\vec{u}\| \cos(0°) = \|\vec{u}\|^2$, where we have said that the angle between $\vec{u}$ and itself is $0°$. It follows that $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$.

Problem 10.

Let $\vec{u}, \vec{v}, \vec{w}$ be vectors, and let α, β be scalars.

True or False: $\vec{u} \cdot (\alpha\vec{v} + \beta\vec{w}) = \alpha\vec{u} \cdot \vec{v} + \beta\vec{u} \cdot \vec{w}$.

Solution: True. This is the "linear" property of the dot product. See Fact 10.

3 Matrices

Problem 11.

Let

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix},$$

and let $\vec{x} = (0, 1, 0, 2, 0)^T$.

What is $A\vec{x}$?

Solution: $(2, 7, 12)^T + (8, 18, 28)^T = (10, 25, 40)^T$.

Rather than carrying out the full matrix-vector multiplication, you might want to instead use Fact 12, which says that $A\vec{x}$ is a linear combination of the columns of A, where the coefficients are the entries of $\vec{x}$. So here, the result should be one copy of the second column of A and two copies of the fourth column of A.

Problem 12.

Let A, B, C, X be matrices of appropriate dimensions. True or False: $X(AB+C)^T = XB^TA^T + XC^T$.

Solution: True. Matrix-matrix multiplication is distributive (see Fact 16).

Also, the transpose of a product is the product of the transposes in reverse order. That is, $(AB)^T = B^TA^T$. See Fact 17.

Problem 13.

Let A, B and C be matrices of appropriate dimensions.

True or False: $ABC = CBA$.

Solution: False. Matrix multiplication is not commutative. See Fact 16.

4 What type of object?

Problem 14.

Let $\vec{x} \in \mathbb{R}^d$ and let A be a $d \times d$ matrix. What type of object is $\vec{x}^T A \vec{x}$?

Solution: A scalar. See Fact 20.

Namely, $A\vec{x}$ is a d-dimensional vector, which we can think of as a $d \times 1$ matrix. Then $\vec{x}^T A \vec{x}$ is a $1 \times d$ matrix times a $d \times 1$ matrix, which is a 1×1 matrix, or a scalar.

Problem 15.

Let $\vec{x}^{(1)}, \ldots, \vec{x}^{(n)}$ be d-dimensional vectors. What type of object is:

$$\frac{1}{n}\sum_{i=1}^{n} \vec{x}^{(i)}(\vec{x}^{(i)})^T$$

Solution: A $d \times d$ matrix. See Fact 20.

$\vec{x}^{(i)}$ is a column vector, which we can think of as a $d \times 1$ matrix. On the other hand, $(\vec{x}^{(i)})^T$ is a row vector, which we can think of as a $1 \times d$ matrix. So the product $\vec{x}^{(i)}(\vec{x}^{(i)})^T$ is $(d \times 1)(1 \times d)$, which is a $d \times d$ matrix.

So each term in the sum is a $d \times d$ matrix, and the sum of n such terms is also a $d \times d$ matrix. Dividing by n doesn't change the type of object.