Lab Solutions 03

1. Consider the vectors:

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad u = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}. \tag{3.1}$$

Compute the values (a) $2 \cdot u$, **(b)** $u + v$, **(c) the dot product** $u \cdot v$, **and (d) the squared norm** $||v||_2^2$**.**

The values are given as follows:

$$2 \cdot u = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix} \tag{3.2}$$

$$u + v = \begin{bmatrix} 3 \\ 6 \\ 11 \end{bmatrix} \tag{3.3}$$

$$u \cdot v = 1 \cdot 2 + 2 \cdot 4 + 3 \cdot 8 = 34 \tag{3.4}$$

$$||v||_2^2 = 1^2 + 2^2 + 3^2 = 14. \tag{3.5}$$

They are a straightforward application of the definitions in the notes.

2. Take the function $f : \mathbb{R}^3 \to \mathbb{R}$ **defined by:**

$$f(x, y, z) = x^2 + x \cdot y + sin(y) \cdot z. \tag{3.6}$$

Write down the function for ∇f**.**

For this, we need to do the computation component-wise. We have:

$$\frac{\partial f}{\partial x} = 2x + y \tag{3.7}$$

$$\frac{\partial f}{\partial y} = y + cos(y) \cdot z \tag{3.8}$$

$$\frac{\partial f}{\partial z} = sin(y) \tag{3.9}$$

And the gradient comes from just taking these and putting them into a single vector:

$$\nabla f = \begin{bmatrix} 2x + y \\ y + cos(y) \cdot z \\ sin(y) \end{bmatrix} \tag{3.10}$$

3. Take the matrix A defined as:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix} \tag{3.11}$$

What is the dimension of the output Av using v from question 1? Compute the value Av.

The output will be a 2-dimensional vector. The value is given by:

$$Av = \begin{bmatrix} 1 \cdot 1 + 1 \cdot 2 + 2 \cdot 3 \\ 3 \cdot 1 + 0 \cdot 2 + 1 \cdot 3 \end{bmatrix} \tag{3.12}$$

$$= \begin{bmatrix} 9 \\ 6 \end{bmatrix}. \tag{3.13}$$

4. Take the matrix B defined as:

$$B = \begin{bmatrix} 2 & 3 \\ 7 & 1 \end{bmatrix}. \tag{3.14}$$

Compute the value of $C = B \cdot A$.

$$C = \begin{bmatrix} 11 & 2 & 7 \\ 10 & 7 & 15 \end{bmatrix} \tag{3.15}$$

5. Let $v \in \mathbb{R}^n$ be a vector of dimension n. We can actually view this vector as a matrix with 1 column and n rows. Similarly, v^t is a vector with n columns and 1 row. What is the value of $v^t v$ in terms of the components v_i? Can you write this in terms of the Euclidean-norm?

The value of $v^t v$ is equal to the squared Euclidean-norm:

$$v^t v = \sum_i v_i \cdot v_i = \sum_i v_i^2 \tag{3.16}$$

$$= \left(\sqrt{\sum_i v_i^2} \right)^2 = ||v||_2^2. \tag{3.17}$$

6. Consider the function g defined by:

$$g(b) = b^t x. \tag{3.18}$$

For $b, x \in \mathbb{R}^n$. Note that we consider b to be variable but x to be fixed. Compute the partial derivative

$$\frac{\partial g}{\partial b_k}. \tag{3.19}$$

The partial derivative is given by:

$$\frac{\partial g}{\partial b_k} = \frac{\partial}{\partial b_k} b^t x \tag{3.20}$$

$$= \frac{\partial}{\partial b_k} \sum_i b_i \cdot x_i \tag{3.21}$$

$$= x_k. \tag{3.22}$$

7. Write the gradient $\nabla_b g$ (the subscript just reinforces that this is a function of b and not x) in a compact format. In other words, do not just write the gradient component-wise.

The full gradient, then, is just given by the vector x:

$$\nabla_b(b^t x) = x. \tag{3.23}$$

One way to think of this is to consider the scalar case where the vector is just of dimension one. In this case, it is the equivalent of taking the derivative of a linear function.

8. Let U be a square matrix such that $Q^t = Q^{-1}$. We can any matrix with this property an *orthogonal* matrix. An orthogonal matrix corresponds to a linear map that is simply a rotation of an n-dimensional space, a property that we will try to get some insight on here. Show that the Euclidean-norm is unchanged when applying Q. In other words, show that:

$$||v||_2^2 = ||Qv||_2^2. \tag{3.24}$$

Then, show that the dot product between Qv and Qu is the same as the dot product between u and v. Use these two properties to argue that the action of Q appears to behave like a rotation in n-dimensional space.

The first result comes from:

$$||Qv||_2^2 = v^t Q^t Q v \tag{3.25}$$
$$= v^t I_n v \tag{3.26}$$
$$= v^t v = ||v||_2^2. \tag{3.27}$$

The second is similar, with:

$$(Qv) \cdot (Qu) = (Qv)^t Qu \tag{3.28}$$
$$= v^t Q^t Q u \tag{3.29}$$
$$= v^t u = v \cdot u. \tag{3.30}$$

Both of these derivations are straightforward, but hopefully help you practice and see the benefit of the matrix form of the dot product.

9. Take a square matrix A such that $A^t = A$ and define the function g where:

$$g(b) = b^t A b. \tag{3.31}$$

We will assume that A is fixed and only b is variable. Convince yourself that:

$$g(b) = \sum_i \sum_j a_{i,j} \cdot b_i \cdot b_j. \tag{3.32}$$

This comes from just applying the definition of the matrix formula twice.

10. Now, compute the partial derivative

$$\frac{\partial g}{\partial b_k}. \tag{3.33}$$

Note that is quite a bit more difficult that the other gradient questions I asked so please be careful. Also note that you can look ahead one question to check your answer.

Start by writing $b^t A b$ by isolating the terms that depend on b_k:

$$b^t A b = \sum_i \sum_j a_{i,j} \cdot b_i \cdot b_j \tag{3.34}$$

$$= a_{kk} b_k^2 + \sum_{i \neq k} a_{ik} b_i b_k + \sum_{j \neq k} a_{kj} b_k b_j + \text{(other terms w/o } b_k\text{)}. \tag{3.35}$$

Now, it is much easier to apply the partial derivative because we have split apart the quadratic, linear and constant terms:

$$\frac{\partial}{\partial b_k} b^t Ab = \frac{\partial}{\partial b_k}(a_{kk}b_k^2) + \frac{\partial}{\partial b_k}(\sum_{i\neq k} a_{ik}b_ib_k) + \frac{\partial}{\partial b_k}(\sum_{j\neq k} a_{kj}b_kb_j) \qquad (3.36)$$

$$= 2a_{kk} \cdot b_k + \sum_{i\neq k} a_{ik}b_i + \sum_{j\neq k} a_{kj}b_j \qquad (3.37)$$

I am going to simplify the equation here by grouping one $a_{kk} \cdot b_k$ with the first sum and another one with the second sum. You may not have actually done this until you go the next two questions:

$$\frac{\partial}{\partial b_k} b^t Ab = \sum_i a_{ik}b_i + \sum_j a_{kj}b_j \qquad (3.38)$$

$$(3.39)$$

11. Finally, take the answer to your last question the prove that

$$\nabla_b g(b) = \nabla_b \left(b^t Ab\right) = 2Ab. \qquad (3.40)$$

Using the simplified form of the partial derivatives, the result follows by noticing that we already have the result for the next equation:

$$\nabla_b g(b) = \nabla_b \left(b^t Ab\right) = Ab + A^t b. \qquad (3.41)$$

The result for this question follows by setting $A = A^t$.

(12. Extra) If we remove the assumption that A is symmetric (i.e., $A = A^t$) then the more general form is given by:

$$\nabla_b g(b) = \nabla_b \left(b^t Ab\right) = Ab + A^t b. \qquad (3.42)$$

Try to prove this more general form as well.

I already showed this in the previous question. I split this question as optional because I thought it may be tricky to figure out that you need to add one copy of $a_{kk} \cdot b_k$ to one sum and another copy in the other sum.