MATH 4931 - MSSC 5931 Homework 1

1. **Representing linear functions as matrix multiplication.** Suppose that $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear. Show that there is a matrix $A \in \mathbb{R}^{m\times n}$ such that for all $x \in \mathbb{R}^n$, $f(x) = Ax$. (Explicitly describe how you get the coefficients A_{ij} from f, and then verify that $f(x) = Ax$ for any $x \in \mathbb{R}^n$.) Is the matrix A that represents f unique? In other words, if $\tilde{A} \in \mathbb{R}^{m\times n}$ is another matrix such that $f(x) = \tilde{A}x$ for all $x \in \mathbb{R}^n$, then do we have $\tilde{A} = A$? Either show that this is so, or give an explicit counterexample.

2. **Matrix representation of polynomial differentiation.** We can represent a polynomial of degree less than n,

$$p(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0,$$

as the vector $(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{R}^n$. Consider the linear transformation $\mathcal{D}$ that differentiates polynomials, *i.e.*, $\mathcal{D}p = dp/dx$. Find the matrix D that represents $\mathcal{D}$ (*i.e.*, if the coefficients of p are given by a, then the coefficients of dp/dx are given by Da).

3. **Counting paths in an undirected graph.** Consider an undirected graph with n nodes, and no self loops (*i.e.*, all branches connect two different nodes). Let $A \in \mathbf{R}^{n\times n}$ be the *node adjacency matrix*, defined as

$$A_{ij} = \begin{matrix} 1 & \text{if there is a branch from node } i \text{ to node } j \\ 0 & \text{if there is no branch from node } i \text{ to node } j \end{matrix}$$

Note that $A = A^{\mathsf{T}}$, and $A_{ii} = 0$ since there are no self loops. We can intrepret A_{ij} (which is either zero or one) as the number of branches that connect node i to node j. Let $B = A^k$, where $k \in \mathbb{Z}$, $k \geq 1$. Give a simple interpretation of B_{ij} in terms of the original graph. (You might need to use the concept of a *path* of length m from node p to node q.)

4. **Gradient of some common functions.** Recall that the gradient of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, at a point $x \in \mathbb{R}^n$, is defined as the vector

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix},$$

where the partial derivatives are evaluated at the point x. The first order Taylor approximation of f, near x, is given by

$$\hat{f}_{\text{tay}}(z) = f(x) + \nabla f(x)^{\mathsf{T}}(z - x).$$

This function is affine, *i.e.*, a linear function plus a constant. For z near x, the Taylor approximation $\hat{f}_{\text{tay}}$ is very near f. Find the gradient of the following functions. Express the gradients using matrix notation.

a) $f(x) = a^{\mathsf{T}}x + b$, where $a \in \mathbb{R}^n$, $b \in \mathbb{R}$.

b) $f(x) = x^{\mathsf{T}}Ax$, for $A \in \mathbb{R}^{n \times n}$.

c) $f(x) = x^{\mathsf{T}}Ax$, where $A = A^{\mathsf{T}} \in \mathbb{R}^{n \times n}$. (Yes, this is a special case of the previous one.)

5. Express the following statements in matrix language. You can assume that all matrices mentioned have appropriate dimensions. Here is an example: "Every column of C is a linear combination of the columns of B" can be expressed as "$C = BF$ for some matrix F".

There can be several answers; one is good enough for us.

a) Suppose Z has n columns. For each i, row i of Z is a linear combination of rows $i, \dots, n$ of Y.

b) W is obtained from V by permuting adjacent odd and even columns (*i.e.*, 1 and 2, 3 and 4, ...).

c) Each column of P makes an acute angle with each column of Q.

d) Each column of P makes an acute angle with the corresponding column of Q.

e) The first k columns of A are orthogonal to the remaining columns of A.