CS365 Written Assignment 1

Khoa Cao

Due

Question 1

1. Let $\mathbf{X} = \mathbf{AB}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Let $\vec{a_i}$ be the i^{th} row of $\mathbf{A}$ and $\vec{b_j}$ be the j^{th} column of $\mathbf{B}$. Then, the ij^{th} entry of $\mathbf{X}$ is given by

$$X_{ij} = \vec{a_i} \cdot \vec{b_j} \qquad \text{definition of matrix multiplication}$$

$$X_{ij} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj} \qquad \text{definition of dot product}$$

Differentiating w.r.t. A_{lp} gives:

$$\frac{\partial X_{ij}}{\partial A_{lp}} = \frac{\partial(\sum_{k=1}^{n} A_{ik} \cdot B_{kj})}{\partial A_{lp}} \qquad \text{from previous result}$$

2. TODO

3. First, we find the scalar equation for the L_p norm of $\vec{x}$.

$$\begin{aligned} ||\vec{x}||_p^p &= \left[\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}\right]^p && \text{definition of } L_p \text{ norm} \\ &= \sum_{i=1}^{n} |x_i|^p && \text{simplifying} \end{aligned}$$

Then, differentiating w.r.t. x_i gives:

$$\begin{aligned} \frac{\partial ||\vec{x}||_p^p}{\partial x_i} &= p|x_i|^{p-1} \cdot \text{sign}(x_i) && \text{chain rule} \\ \Rightarrow \frac{\partial ||\vec{x}||_p}{\partial \vec{x}} &= \begin{bmatrix} \frac{\partial ||\vec{x}||_p^p}{\partial x_1} & \frac{\partial ||\vec{x}||_p^p}{\partial x_2} & \cdots & \frac{\partial ||\vec{x}||_p^p}{\partial x_n} \end{bmatrix} && \text{writing derivative in vector form} \\ &= \begin{bmatrix} p|x_1|^{p-1} \cdot sgn(x_1) & \cdots & p|x_n|^{p-1} \cdot sgn(x_n) \end{bmatrix} && \text{substituting previous result} \end{aligned}$$

4. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. First, we find the scalar equation for $\vec{x}^T \mathbf{A} \vec{x}$.

$$\begin{aligned}
\vec{x}^T \mathbf{A} \vec{x} &= \begin{bmatrix} \vec{x} \cdot \vec{a_1} & \vec{x} \cdot \vec{a_2} & \cdots & \vec{x} \cdot \vec{a_n} \end{bmatrix} \cdot \vec{x} && \text{definition of matrix multiplication} \\
&= \sum_{i=1}^{n} (\vec{x} \cdot \vec{a_i}) x_i && \text{definition of dot product} \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ji} x_j \right) x_i && \text{expanding } \vec{x} \cdot \vec{a_i} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ji} x_j x_i && \text{simplifying}
\end{aligned}$$

Then, differentiating w.r.t. x_k gives:

$$\begin{aligned}
\frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial x_k} &= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ji} \frac{\partial (x_j x_i)}{\partial x_k} && \text{linearity of differentia} \\
&= \sum_{i=1}^{n} A_{ki} x_i + \sum_{j=1}^{n} A_{jk} x_j && \text{only terms with } i = k \text{ or } j = k \text{ re} \\
\Rightarrow \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial \vec{x}} &= \begin{bmatrix} \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial x_1} & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial x_2} & \cdots & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial x_n} \end{bmatrix} && \text{writing derivative in vector} \\
&= \begin{bmatrix} \sum_i A_{1i} x_i + \sum_j A_{j1} x_j & \cdots & \sum_i A_{ni} x_i + \sum_j A_{jn} x_j \end{bmatrix} && \text{substituting previous r} \\
&= \begin{bmatrix} \sum_i A_{1i} x_i & \cdots & \sum_i A_{ni} x_i \end{bmatrix} + \begin{bmatrix} \sum_j A_{j1} x_j & \cdots & \sum_j A_{jn} x_j \end{bmatrix} && \text{separating the s} \\
&= \mathbf{A} \vec{x} + \mathbf{A}^T \vec{x} && \text{rewriting in matrix} \\
&= (\mathbf{A} + \mathbf{A}^T) \vec{x} && \text{factoring o}
\end{aligned}$$

5. Using the result from part (d), we have:

$$\vec{x}^T \mathbf{A} \vec{x} = \sum_{i} \sum_{j} A_{ji} x_j x_i \qquad \text{from part (d)}$$

Differentiating w.r.t. A_{kl} gives:

$$
\begin{aligned}
\frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{kl}} &= \sum_i \sum_j \frac{\partial (A_{ji} x_j x_i)}{\partial A_{kl}} && \text{linearity of differentiation} \\
&= x_l x_k && \text{only the term with } j = k \text{ and } i = l \text{ remains} \\
\Rightarrow \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial \mathbf{A}} &= \begin{bmatrix} \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{11}} & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{12}} & \cdots & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{1n}} \\ \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{21}} & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{22}} & \cdots & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{n1}} & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{n2}} & \cdots & \frac{\partial \vec{x}^T \mathbf{A} \vec{x}}{\partial A_{nn}} \end{bmatrix} && \text{writing derivative in matrix form} \\
&= \begin{bmatrix} x_1 x_1 & x_1 x_2 & \cdots & x_1 x_n \\ x_2 x_1 & x_2 x_2 & \cdots & x_2 x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n x_1 & x_n x_2 & \cdots & x_n x_n \end{bmatrix} && \text{substituting previous result} \\
&= \vec{x} \vec{x}^T && \text{rewriting in vector form}
\end{aligned}
$$

Question 3

1. Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Suppose $\mathbf{M}$ has eigenvectors $\vec{v_1}, \vec{v_2}$ with distinct eigenvalues λ_1, λ_2. We will show that these eigenvectors are orthogonal, i.e. $\vec{v_1} \cdot \vec{v_2} = 0$. Consider:

$$\begin{aligned}
\mathbf{M}\vec{v_1} \cdot \vec{v_2} &= \vec{v_1} \cdot \mathbf{M}^T\vec{v_2} && \text{given lemma} \\
\Rightarrow \mathbf{M}\vec{v_1} \cdot \vec{v_2} &= \vec{v_1} \cdot \mathbf{M}\vec{v_2} && \text{since } \mathbf{M} \text{ is symmetric} \\
\Rightarrow \lambda_1\vec{v_1} \cdot \vec{v_2} &= \lambda_2\vec{v_1} \cdot \vec{v_2} && \text{definition of eigenvector} \\
\Rightarrow \lambda_1\vec{v_1} \cdot \vec{v_2} - \lambda_2\vec{v_1} \cdot \vec{v_2} &= 0 && \text{rearranging} \\
\Rightarrow (\lambda_1 - \lambda_2)(\vec{v_1} \cdot \vec{v_2}) &= 0 && \text{factoring} \\
\Rightarrow \vec{v_1} \cdot \vec{v_2} &= 0 && \text{since } \lambda_1 \neq \lambda_2
\end{aligned}$$

   Therefore, eigenvectors of a square symmetric matrix with distinct eigenvalues are orthogonal.

2. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be any matrix. We will show that $\mathbf{A}\mathbf{A}^T \in \mathbb{R}^{m\times m}$ and $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n\times n}$ are symmetric. Consider:

$$\begin{aligned}
(\mathbf{A}\mathbf{A}^T)^T &= (\mathbf{A}^T)^T\mathbf{A}^T && \text{transpose of a matrix product} \\
&= \mathbf{A}\mathbf{A}^T && \text{double transpose} \\
&\Rightarrow \mathbf{A}\mathbf{A}^T \text{ is symmetric by definition}
\end{aligned}$$

   Similarly, consider:

$$\begin{aligned}
(\mathbf{A}^T\mathbf{A})^T &= \mathbf{A}^T(\mathbf{A})^T && \text{transpose of a matrix product} \\
&= \mathbf{A}^T\mathbf{A} && \text{double transpose} \\
&\Rightarrow \mathbf{A}^T\mathbf{A} \text{ is symmetric by definition}
\end{aligned}$$