Homework 2.

Question 1.

(a) Let x be an n dimensional vector, and M be an $n \times n$ dimensional matrix. Show that

$$x^T M x = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} x_i x_j.$$

Solution:

$$\begin{aligned}
(Mx)_i &= \sum_{j=1}^{n} m_{ij} x_j, \quad i = 1, 2, \ldots, n \\
x^T &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}, \\
Mx &= \begin{bmatrix} (Mx)_1 \\ (Mx)_2 \\ \vdots \\ (Mx)_n \end{bmatrix} \\
x^T(Mx) &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} (Mx)_1 \\ (Mx)_2 \\ \vdots \\ (Mx)_n \end{bmatrix} \\
&= x_1 (Mx)_1 + x_2 (Mx)_2 + \cdots + x_n (Mx)_n \\
&= \sum_{i=1}^{n} x_i (Mx)_i
\end{aligned}$$

$$
\begin{aligned}
x^T(Mx) &= x_1(Mx)_1 + x_2(Mx)_2 + \cdots + x_n(Mx)_n \\
&= x_1\left(\sum_{j=1}^{n} m_{1j}x_j\right) + x_2\left(\sum_{j=1}^{n} m_{2j}x_j\right) + \cdots + x_n\left(\sum_{j=1}^{n} m_{nj}x_j\right) \\
&= \big(m_{11}x_1x_1 + m_{12}x_1x_2 + \cdots + m_{1n}x_1x_n\big) \\
&\quad + \big(m_{21}x_2x_1 + m_{22}x_2x_2 + \cdots + m_{2n}x_2x_n\big) \\
&\quad + \cdots \\
&\quad + \big(m_{n1}x_nx_1 + m_{n2}x_nx_2 + \cdots + m_{nn}x_nx_n\big) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}x_ix_j
\end{aligned}
$$

(b) Let b and x be vectors in $\mathbb{R}^n$. Show that

Solution:

$$\nabla(b^T x) = b.$$

$$\begin{aligned} b^T x &= \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \end{aligned}$$

$$\frac{\partial}{\partial x_1}(b^T x) = b_1, \quad \frac{\partial}{\partial x_2}(b^T x) = b_2, \quad \ldots, \quad \frac{\partial}{\partial x_n}(b^T x) = b_n$$

$$\nabla_x(b^T x) = \begin{bmatrix} \frac{\partial}{\partial x_1}(b^T x) \\ \frac{\partial}{\partial x_2}(b^T x) \\ \vdots \\ \frac{\partial}{\partial x_n}(b^T x) \end{bmatrix}$$

$$\nabla_x(b^T x) = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b$$

(c) Let A be an $n \times n$ matrix and $x \in \mathbb{R}^n$. Show that

$$\nabla(x^T A x) = (A + A^T)x.$$

Solution:

$$f(x) = x^T A x \;=\; \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i x_j$$

$$\begin{aligned}
\frac{\partial f}{\partial x_1} &= \frac{\partial}{\partial x_1} \Big(\sum_{j=1}^{n} a_{1j}\, x_1 x_j + \sum_{i=2}^{n} \sum_{j=2}^{n} a_{ij}\, x_i x_j + \sum_{i=1}^{n} a_{i1}\, x_i x_1 \Big) \\
&= \sum_{j=1}^{n} a_{1j}\, x_j \;+\; \sum_{i=1}^{n} a_{i1}\, x_i \\
&= (Ax)_1 \;+\; (A^T x)_1.
\end{aligned}$$

$$\frac{\partial f}{\partial x_k} = (Ax)_k + (A^T x)_k$$

$$\nabla_x \big(x^T A x \big) = Ax + A^T x = (A + A^T) x$$

(d) Let $x \in \mathbb{R}^n$ and

$$f(x) = x^T Ax + b^T x + c$$

where A is a symmetric $n \times n$ matrix, b is an n dimensional vector and c is a scalar. Assume A is invertible and solve for the critical point of $f(x)$.

Solution:

$$\begin{aligned} f(x) &= x^T Ax + b^T x + c \\ \nabla f(x) &= 2Ax + b = 0 \\ 2Ax &= -b \\ Ax &= -\tfrac{1}{2}b \\ x &= -\tfrac{1}{2}A^{-1}b \end{aligned}$$

(e) Let $v^{(1)}, v^{(2)}, \ldots, v^{(k)} \in \mathbb{R}^n$ be linearly independent vectors in $\mathbb{R}^n$.
Let $x \in \mathbb{R}^n$.
Consider the projection of x onto

$$\Omega = \text{span}(v^{(1)}, v^{(2)}, \ldots, v^{(k)}).$$

Solution:

Working through the 2-D case but this solution works for the n-D case well.

$$\begin{aligned} L(\alpha) &= \|x - V\alpha\|^2 = (x - V\alpha)^T(x - V\alpha) \\ &= x^T x - 2\,\alpha^T V^T x + \alpha^T V^T V\alpha. \end{aligned}$$

$$\nabla_\alpha L(\alpha) = -2V^T x + 2V^T V\,\alpha = 0$$

$$\Rightarrow \quad V^T V\,\alpha = V^T x$$

$$\Rightarrow \quad \alpha = (V^T V)^{-1} V^T x \qquad \text{(since the columns of } V \text{ are linearly independent).}$$

$$x_p = V\alpha \;=\; V\,(V^T V)^{-1} V^T x.$$

Question 2.

(a) Show that $Q^{-1} = Q^T$.

Solution:

$$\begin{aligned} Q^T Q &= I \qquad \text{(by definition of orthonormal)} \\ Q^{-1} Q &= I \qquad \text{(by definition of inverse)} \\ \therefore \quad Q^T &= Q^{-1} \end{aligned}$$

(b) Show that $\|Qx\| = \|x\|$ for all $x \in \mathbb{R}^n$ if Q is an $n \times n$ orthonormal matrix.(Hint: Show that $\|Qx\|^2 = x^T Q^T Q x$.)

Solution:

$$\begin{aligned} \|Qx\|^2 &= (Qx)^T (Qx) \quad \text{(by definition of norm)} \\ &= x^T Q^T Q x \\ &= x^T Q^{-1} Q x \\ &= x^T I x \\ &= x^T x \\ &= \|x\|^2 \\ \therefore \quad \|Qx\| &= \|x\| \end{aligned}$$

(c) Suppose Q is a 2×2 orthonormal matrix. Show that Q has the following form,

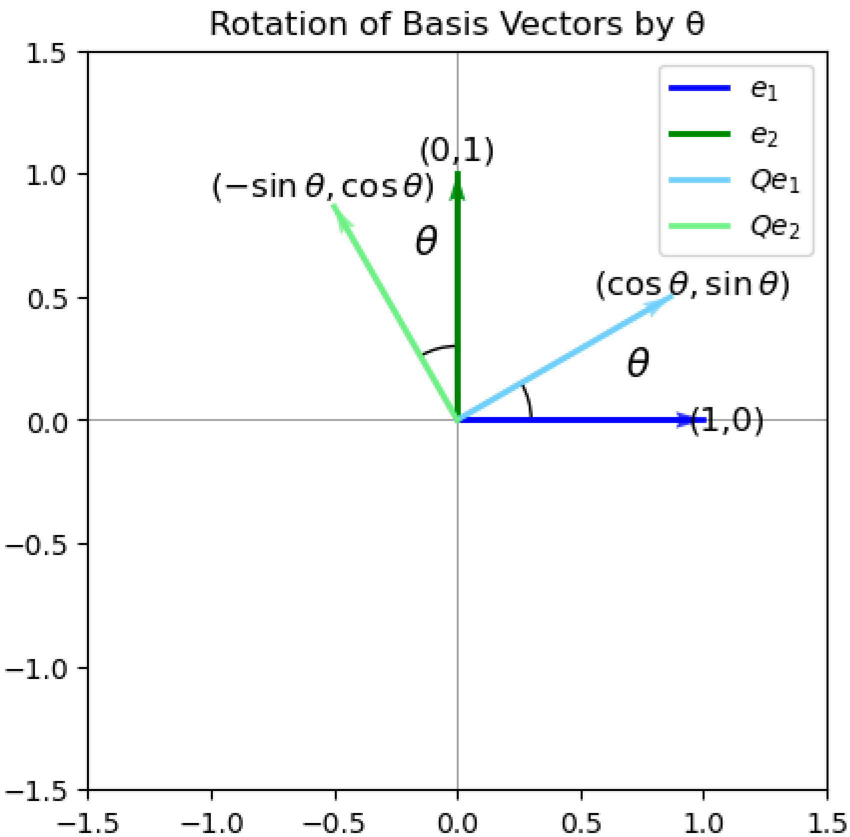
Solution:

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad Q = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Show that for the first form of Q, Qx is a rotation of x by θ degrees. You can do this for general x or by considering Qx for $x = (1, 0)$ and $x = (0, 1)$. (The second form of Q is a reflection about the y-axis followed by a rotation.)

$$\text{let } e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{aligned} Q &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ Qe_1 &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\ Qe_2 &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \end{aligned}$$



Rotation of Basis Vectors by θ
e_1
e_2
Qe_1
Qe_2
(0,1)
$(-\sin\theta, \cos\theta)$
θ
$(\cos\theta, \sin\theta)$
θ
(1,0)
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

Using a specific example of $\theta = \frac{\pi}{6}$

$$Q(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$Q(\theta)e_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad Q(\theta)e_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$\text{For } \theta = 30^\circ = \tfrac{\pi}{6}:$$

$$Q\left(\tfrac{\pi}{6}\right) = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

$$Q\left(\tfrac{\pi}{6}\right)e_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}, \quad Q\left(\tfrac{\pi}{6}\right)e_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}.$$

Using a general θ

$$\|e_1\|^2 = \|(1,0)\|^2 = 1, \qquad \|Qe_1\|^2 = \|(\cos\theta,\ \sin\theta)\|^2 = \cos^2\theta + \sin^2\theta = 1.$$

$$\|e_2\|^2 = \|(0,1)\|^2 = 1, \qquad \|Qe_2\|^2 = \|(-\sin\theta,\ \cos\theta)\|^2 = \sin^2\theta + \cos^2\theta = 1.$$

$$\therefore \quad \|Qe_1\| = \|e_1\| = 1, \qquad \|Qe_2\| = \|e_2\| = 1.$$

$$\begin{aligned}(Qe_1)\cdot(Qe_2) &= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \cdot \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \\ &= \cos\theta(-\sin\theta) + \sin\theta\cos\theta \\ &= 0\end{aligned}$$

As we see above, the matrix Q has not changed the magnitude of the vectors, (1,0) and (0,1), instead only rotated it theta degrees and that the column vectors of Q are still orthogonal - showing that Q is still orthonormal.

Question 3.

Let $\Sigma \in \mathbb{R}^{n \times n}$ be a symmetric matrix.
Show that there exists an orthonormal matrix Q and a diagonal matrix D such that

$$\Sigma = QDQ^T,$$

where the columns of Q are the eigenvectors of Σ and the diagonal entries of D are the corresponding eigenvalues.

Solution:

$$\text{Let } Q = \begin{bmatrix} q^{(1)} & \cdots & q^{(n)} \end{bmatrix}, \;\; D = \text{diag}(\lambda_1, \dots, \lambda_n),$$
$\Sigma q^{(i)} = \lambda_i q^{(i)}, \;\; Q^T Q = I.$

$$\begin{aligned} Q^T q^{(i)} &= e_i \qquad \text{where } e_i \text{ is the } i\text{-th standard basis vector} \\ QDQ^T q^{(i)} &= QD(Q^T q^{(i)}) = QDe_i = Q(\lambda_i e_i) = \lambda_i \, Qe_i = \lambda_i \, q^{(i)} \\ &= \Sigma q^{(i)} \qquad \text{(by the eigenpair definition)} \end{aligned}$$

$$\Rightarrow \quad \Sigma = QDQ^T \quad \text{(since they agree on the orthonormal basis } \{q^{(i)}\}\text{)}.$$