1) a) $f(x) = \frac{1}{2} x^T A x + b^T x$; calc. $\nabla f(x)$ when A is symmetric (i.e. $A^T = A$)

$\nabla f(x) \Rightarrow$ in this case is called "Gradient of the Quadratic Form."

Solution from lectures:

Assuming we know

i) $x^T A x = Tr(x^T A x)$

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

iii) $A^T = A \longrightarrow (A + A^T)x = 2Ax$

$$\nabla_x\left(\frac{1}{2} x^T A x\right) = \frac{1}{2} \times 2Ax = Ax$$

Now, $\frac{\partial(x^T y)}{\partial x_i} = y$

$$\therefore \frac{\partial(b^T x)}{\partial x_i} = \frac{\partial(x^T b)}{\partial x_i} = b$$

$$\therefore \nabla f(x) = Ax + b$$

Otherwise:

$$\frac{\partial A^T x}{\partial x} = \frac{\partial x^T A}{\partial x} = A^T \quad ; \quad \frac{\partial x}{\partial x} = I$$

Identity matrix

$$\frac{df(g(x), h(x))}{dx} = g^T(x) \frac{\partial h(x)}{\partial x} + h^T(x) \frac{\partial g(x)}{\partial x}$$

$$= x^T \frac{\partial Ax}{\partial x} + (Ax)^T \frac{\partial x}{\partial x}$$

$$= x^T A + x^T A^T I$$

$$= x^T (A + A^T)$$

$$\nabla f = \left(\frac{df}{dx}\right)^T = \left(x^T (A + A^T)\right)^T$$
$$= (A^T + A)x$$

SOLUTIONS

b) $f(x) = g(h(x))$ $\quad g : \mathbb{R} \to \mathbb{R}$
$\quad h : \mathbb{R}^n \to \mathbb{R}$

$$\frac{\partial g(h(x))}{\partial x_i} = \frac{\partial g(h(x))}{\partial (\cancel{h}(x))} \cdot \frac{\partial (h(x))}{\partial x_i}$$
$$= g' h(x) \cdot \frac{\partial (h(x))}{\partial x_i}$$
$$= g' h(x) \nabla h(x)$$

$$\therefore \nabla f(x) = \nabla g(h(x))$$
$$= g'(h(x)) \nabla h(x)$$

c) $f(x) = \frac{1}{2} x^T A x + b^T x$; A is symmetric
$b \in \mathbb{R}^n$ is a vector. $\nabla^2 f(x)$? (find hessian)

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \frac{\partial^2}{\partial x_1 \partial x_2} f(x) & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2}{\partial x_2^2} f(x) & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2}{\partial x_n \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_n^2} f(x) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial \nabla f(x)}{\partial x_1} & \frac{\partial \nabla f(x)}{\partial x_2} & \cdots & \frac{\partial \nabla f(x)}{\partial x_n} \end{bmatrix}$$

we know $\nabla f(x) = Ax + b$

$$= \begin{bmatrix} \frac{\partial (Ax+b)}{\partial x_1} & \frac{\partial (Ax+b)}{\partial x_2} & \cdots \end{bmatrix}$$

$$= \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} = A$$

d) i) $f(x) = g(a^T x)$ $\quad g: \mathbb{R} \to \mathbb{R}$ is differentiable
$a \in \mathbb{R}^n$ is a vector

$$= \frac{\partial g(a^T x)}{\partial (a^T x)} \cdot \frac{\partial (a^T x)}{\partial x_i}$$

$$= g'(a^T x)\, a$$

ii) We may generalize the hessian to

$$\frac{\partial^2 f}{\partial x_i \partial x_j}$$

$$\frac{\partial^2 (g(h(x)))}{\partial x_i \partial x_j} = \frac{\partial^2 g(h(x))}{\partial (h(x))^2} \frac{\partial (h(x))}{\partial x_i} \frac{\partial (h(x))}{\partial x_j}$$

$$= g''(h(x)) \frac{\partial h(x)}{\partial x_i} \frac{\partial h(x)}{\partial x_j}$$

SOLUTIONS

Now $$\frac{\partial^2 g(a^T x)}{\partial x_i \partial x_j} = g''(a^T x) \frac{\partial (a^T x)}{\partial x_i} \frac{\partial (a^T x)}{\partial x_j}$$

We know, $$\frac{\partial (a^T x)}{\partial x_i} = a_i$$

$$\therefore \quad g''(a^T x) \frac{\partial (a^T x)}{\partial x_i} \frac{\partial (a^T x)}{\partial x_j} = g''(a^T x) a_i a_j$$

$$\nabla^2 f(x) = \nabla^2 g(a^T x)$$

$$= g'' a^T x \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= g'' a^T x \, a a^T$$

2) a) $z \in \mathbb{R}^n$ be an n-vector

(3)

$A = zz^T$; prove A is positive semidefinite

$$A = zz^T$$

$$A^T = (zz^T)^T = z^T (z^T)^T = z^T z$$

if z is invertible then, take any $x \in \mathbb{R}^n$

$$x^T z z^T x = (z^T x)^T (z^T x) > 0$$

$$= (x^T z)^2 > 0$$

b) If $n=1$, the null space is a zero/null vector and it's dimension is zero.

$\{0\}$ has dimesion 0

A vector in itself has no dimension

A subspace has dimension

$\mathbb{R}^3$ has dimension 3 because we find it in a linearly independent set with 3 elements

Subsets of $\{0\}$ is empty set & whole set

In whole set zero vector is linearly dependent (you can't find a zero linear combination with non-zero coefficients out of its elements), so only linearly independent set is empty set which has dimesion 0.

c) $A \in \mathbb{R}^{n \times n} \longrightarrow$ positive semidefinite

$B \in \mathbb{R}^{m \times n} \longrightarrow$ arbitrary where $m, n \in \mathbb{N}$

is BAB^T a positive semidefinite? Yes

for $x \in \mathbb{R}^m$ we can define a vector $V = B^T x \in \mathbb{R}$

$$x^T (BAB^T) x = (B^T x)^T A (B^T x)$$

$= v^T A v \geq 0$

as this $(v^T A v)$ holds true for any vector v.

3) a) $A \in \mathbb{R}^{n \times n}$ is diagonizable i.e. $A = T \Lambda T^{-1}$

where T is an invertible $T \in \mathbb{R}^{n \times n}$

$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2 \cdots \lambda_n)$ is diagonal

$t^{(i)}$ denote the cols of T, $T = [t^{(1)}, t^{(2)} \ldots t^{(n)}]$

$t^{(i)} \in \mathbb{R}^n$

Show $A t^{(i)} = \lambda_i t^{(i)}$

so that eigenvector pairs ~~of~~ A are $(t^{(i)}, \lambda_i)$

Proof : If T is invertible

$$T^{-1} A T = \Lambda \quad \& \quad A = T \Lambda T^{-1}$$

$$A = T \Lambda T^{-1}$$

$$A T = T \Lambda$$

$$A [t^{(1)}, t^{(2)} \ldots t^{(n)}] = [t^{(1)}, t^{(2)} \ldots t^{(n)}] \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$[A t^{(1)}, A t^{(2)} \ldots A t^{(n)}] = [\lambda_1 t^{(1)}, \lambda_2 t^{(2)} \ldots \lambda_n t^{(n)}]$$

$$A t^{(i)} = \lambda_i t^{(i)}$$

where $\{(\lambda_i, t_i)\}_{i=1}^{n}$ are the eigenvalue-eigenvector pairs of A.

b) $U \in \mathbb{R}^{n \times n}$ is orthogonal if $U^T U = I$ ④

Spectral Theorem

$A \in \mathbb{R}^{n \times n}$ is symmetric ie $A = A^T$ then A is diagonalizable by a real orthogonal matrix

$\Lambda \in \mathbb{R}^{n \times n}$ is a diagonal matrix

$U \in \mathbb{R}^{n \times n}$ is orthogonal such that $U^T A U = \Lambda$ or $A = U \Lambda U^T$

& $\lambda_i = \lambda_i(A)$ is the ith eigenvalue of A

if $U = [u^{(1)} \dots u^n]$ is orthogonal $u^{(i)} \in \mathbb{R}^n$ & $A = U \Lambda U^T$, then $u^{(i)}$ is eigenvector of A & $A u^{(i)} = \lambda_i u^{(i)}$, $\Lambda = \mathrm{diag}(\lambda_1 \dots \lambda_n)$

(just like previous question)

$$A = U \Lambda U^T \quad \text{or} \quad U A U^T = \Lambda$$

$$A U = U \Lambda U^T U$$

$$A U = U \Lambda$$

$$A \left[u^{(1)} \; u^{(2)} \; \dots \; u^{(n)}\right] = \left[u^{(1)} \; u^{(2)} \dots u^{(n)}\right] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ 0 & \lambda_2 & 0 \\ \vdots & & \ddots \\ 0 & 0 & \lambda_n \end{bmatrix}$$

$$\left[A u^{(1)} \;\; A u^{(2)} \dots A u^{(n)}\right] = \left[\lambda_1 u^{(1)} \;\; \lambda_2 u^{(2)} \dots \lambda_n u^{(n)}\right]$$

$$A u^{(i)} = \lambda_i u^{(i)}$$

c) $$A t^{(i)} = \lambda_i t^{(i)}$$

$$(t^{(i)})^T A t^{(i)} = \lambda_i \| t^{(i)} \|_2$$

$$= \lambda_i \geq 0$$