Problem Set 0 Solution

Problem 1

(a)

When $f(x)=\frac{1}{2}\vec{x}^T A\vec{x}+\vec{b}^T\vec{x}$, we can derive that

$$\nabla f(x)=\nabla\left(\frac{1}{2}\vec{x}^T A\vec{x}+\vec{b}^T\vec{x}\right)=A\vec{x}+\vec{b}$$

(b)

When $f(x)=g(h(x))$, we can derive that

$$\begin{aligned}\nabla f(x)&=\frac{\partial}{\partial x_i}f(x)=\frac{\partial}{\partial x_i}g(h(x))=\frac{\partial g(h(x))}{\partial x_i}\\&=\frac{\partial g(h(x))}{\partial h(x)}\frac{\partial h(x)}{\partial x_i}\\&=g'(h(x))\nabla h(x)\end{aligned}$$

(c)

When $f(x)=\frac{1}{2}\vec{x}^T A\vec{x}+\vec{b}^T\vec{x}$, we have obtained that

$$\nabla f(x)=A\vec{x}+\vec{b}$$

For the given matrices

$$A=\begin{pmatrix}A_{11}&\cdots&A_{1n}\\\vdots&\ddots&\vdots\\A_{n1}&\cdots&A_{nn}\end{pmatrix},\ \vec{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\text{ and }\vec{b}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

We can derive that

$$A\vec{x}+b=\begin{pmatrix}A_{11}x_1+A_{12}x_2+\cdots+A_{1n}x_n+b_1\\\vdots\\A_{n1}x_1+A_{n2}x_2+\cdots A_{nn}x_n+b_n\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^{n}A_{1i}x_i+b_1\\\vdots\\\sum_{i=1}^{n}A_{ni}x_i+b_n\end{pmatrix}$$

Thus, from an intuitive view, we can derive that

$$\nabla^2 f(x)=\left(\frac{\partial\nabla f(x)}{\partial x_1}\ \ \frac{\partial\nabla f(x)}{\partial x_2}\ \cdots\ \frac{\partial\nabla f(x)}{\partial x_n}\right)=\left(\frac{\partial\nabla\left(A\vec{x}+\vec{b}\right)}{\partial x_1}\ \ \frac{\partial\nabla\left(A\vec{x}+\vec{b}\right)}{\partial x_2}\ \cdots\ \frac{\partial\nabla\left(A\vec{x}+\vec{b}\right)}{\partial x_n}\right)$$

$$= \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} = A$$

(d)

When $f(x) = g(\vec{a}^T \vec{x})$, we can derive that

$$\nabla f(x) = \nabla g(\vec{a}^T \vec{x}) = g'(\vec{a}^T \vec{x}) \nabla (\vec{a}^T \vec{x}) = g'(\vec{a}^T \vec{x}) \vec{a}$$

To derive the second-order partial derivatives, let $f(x) = g(h(x))$ first

Thus, we can derive that

$$\begin{aligned} \nabla^2 f(x) = \nabla^2 g(h(x)) &= \frac{\partial^2 g(h(x))}{\partial x_i x_j} \\ &= \frac{\partial^2 g(h(x))}{\partial^2 (h(x))^2} \frac{\partial h(x)}{\partial x_i} \frac{\partial h(x)}{\partial x_j} \\ &= g''(h(x)) \frac{\partial h(x)}{\partial x_i} \frac{\partial h(x)}{\partial x_j} \end{aligned}$$

Thus, in this case, we can derive that

$$\nabla^2 f(x) = g''(\vec{a}^T \vec{x}) \frac{\partial (\vec{a}^T \vec{x})}{\partial x_i} \frac{\partial (\vec{a}^T \vec{x})}{\partial x_j} = g''(\vec{a}^T \vec{x}) a_i a_j (1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n)$$

To span all the available values for i and j, which is the matrix including all the combinations of $a_i a_j$, thus

$$\nabla^2 f(x) = g''(\vec{a}^T \vec{x}) \begin{pmatrix} a_1 a_1 & a_1 a_2 & \cdots & a_n a_1 \\ a_2 a_1 & a_2 a_2 & \cdots & a_n a_2 \\ \vdots & \vdots & \cdots & \vdots \\ a_n a_1 & a_n a_2 & \cdots & a_n a_n \end{pmatrix} = g''(\vec{a}^T \vec{x}) \vec{a} \vec{a}^T$$

Problem 2

(a) *Proof*

Let $\vec{z}=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$, thus we can obtain that $\vec{z}^T=\begin{pmatrix} z_1 & z_2 & \cdots & z_n \end{pmatrix}$

Thus, we can derive that

$$\vec{z}\vec{z}^T=\begin{pmatrix} z_1z_1 & z_1z_2 & \cdots & z_nz_1 \\ z_2z_1 & z_2z_2 & \cdots & z_nz_2 \\ \vdots & \vdots & \cdots & \vdots \\ z_nz_1 & z_nz_2 & \cdots & z_nz_n \end{pmatrix}$$

Since $(\vec{z}\vec{z}^T)^T=(\vec{z}^T)^T\vec{z}^T=\vec{z}\vec{z}^T$, and for $\vec{x}\in\mathbb{R}^n$,

we have $\vec{x}^TA\vec{x}=\vec{x}^T\vec{z}\vec{z}^T\vec{x}=(\vec{x}^T\vec{z})^2\geqslant 0$ (since $\vec{x}^T\vec{z}=\vec{z}^T\vec{x}$)

Thus, $A=\vec{z}\vec{z}^T$ is positive semidefinite.

(b)

Since the vector $\vec{z}$ is a n-by-1 matrix, and $\vec{x}^T$ is a 1-by-n matrix, we can find that $\vec{x}^T\vec{z}$ is a real number. (1-by-1 matrix)

Thus, the null-space of A is $\mathrm{Nul}(A)=\{x\in\mathbb{R}^n\colon\ \vec{x}^T\vec{z}=0\}$

(c)

BAB^T is positive semidefinite, here is the proof.

Since $(BAB^T)^T=(B^T)^T(BA)^T=BA^TB^T$, and $A^T=A$,

$(BAB^T)^T=BAB^T$

For $\vec{x}\in\mathbb{R}^{n\times n}$, we have $\vec{x}^TBAB^T\vec{x}=(\vec{x}^TB)A(\vec{x}^TB)^T\geqslant 0$

Thus, BAB^T is positive semidefinite.

Problem 3

(a) *Proof*

Since we have $A = T\Lambda T^{-1}$, we can obtain that $AT = T\Lambda T^{-1}T = T\Lambda$.

Also, we have $T = \begin{pmatrix} \vec{t}^{(1)} & \cdots & \vec{t}^{(n)} \end{pmatrix}$, and $\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$,

we can derive that

$$T\Lambda = \begin{pmatrix} \vec{t}^{(1)} & \cdots & \vec{t}^{(n)} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \vec{t}^{(1)} & \lambda_2 \vec{t}^{(2)} & \cdots & \lambda_n \vec{t}^{(n)} \end{pmatrix}$$

Thus, we can know that $At^{(i)} = \lambda_i t^{(i)}$, and the proof completes.

(b) *Proof*

Omitted, since it is identical to the first one.

(c) *Proof*

Since we have known that for an orthogonal matrix $U = (\vec{u}^{(1)} \ \cdots \ \vec{u}^{(n)})$, we can obtain that $\vec{u}^{(i)}$ is an eigenvector for a matrix $A = U\Lambda U^T$. For the eigenvector, we can know that $\|\vec{u}^{(i)}\| = 1$. In terms of the conclusion $A\vec{u}^{(i)} = \lambda_i \vec{u}^{(i)}$, we can obtain that

$$(\vec{u}^{(i)})^T A \vec{u}^{(i)} = (\vec{u}^{(i)})^T \lambda_i \vec{u}^{(i)} = \lambda_i \|\vec{u}^{(i)}\|^2 = \lambda_i$$

Since A is positive semi-definite, thus, for $\vec{u}^{(i)} \in \mathbb{R}^n$, we have $(\vec{u}^{(i)})^T A \vec{u}^{(i)} \geqslant 0$.

Thus, we can conclude that $\lambda_i(A) \geqslant 0$.