Assignment 1 - Optimization Method

薛劭杰 1930026143

1. If $(2,2)^T$ is a convex combination of $(0,0,)^T$, $(1,4)^T$, $(3,1)^T$, then we can get a system with three equations.

$$\sum_{i=1}^{3} \lambda_i = \lambda_1 + \lambda_2 + \lambda_3 = 1 \quad (1)$$
$$0 * \lambda_1 + 1 * \lambda_2 + 3 * \lambda_3 = 2 \quad (2)$$
$$0 * \lambda_1 + 4 * \lambda_2 + 1 * \lambda_3 = 2 \quad (3)$$

Solve this system, we can get $\left\{ \begin{matrix} \lambda_1 = \frac{1}{11} \\ \lambda_2 = \frac{4}{11} \\ \lambda_3 = \frac{6}{11} \end{matrix} \right\}$. Therefore, $(2,2)^T = \frac{1}{11}(0,0,)^T + \frac{4}{11}(1,4)^T + \frac{6}{11}(3,1)^T$.

2. Function $f(x)$ is convex, it satisfies $f(\lambda\bar{x} + (1-\lambda)\hat{x}) \le \lambda f(\bar{x}) + (1-\lambda)f(\hat{x})$
When it comes to $g(x)$, the left-hand side of this inequality can be rewritten as: $left$: $k * f(\lambda\bar{x} + (1-\lambda)\hat{x})$. And the right -hand side of this inequality can be rewritten as: $\lambda k * f(\bar{x}) + (1-\lambda)k * f(\hat{x}) = k * (\lambda f(\bar{x}) + (1-\lambda)f(\hat{x}))$
If $k > 0$, the left-hand side of the inequality must be less than or equal to right.

$$k * f(\lambda\bar{x} + (1-\lambda)\hat{x}) \le k * \left(\lambda f(\bar{x}) + (1-\lambda)f(\hat{x})\right)$$
$$\Downarrow$$
$$g(\lambda\bar{x} + (1-\lambda)\hat{x}) \le \lambda g(\bar{x}) + (1-\lambda)g(\hat{x})$$
$$\Downarrow$$
$$\text{Function } g(x) \text{ is convex}$$

If $k < 0$, the left-hand side of the inequality must be larger than or equal to right.

$$k * f(\lambda\bar{x} + (1-\lambda)\hat{x}) \ge k * \left(\lambda f(\bar{x}) + (1-\lambda)f(\hat{x})\right)$$
$$\Downarrow$$
$$g(\lambda\bar{x} + (1-\lambda)\hat{x}) \ge \lambda g(\bar{x}) + (1-\lambda)g(\hat{x})$$
$$\Downarrow$$
$$\text{Function } g(x) \text{ is concave}$$

3. We know that $g(x)$ is a concave function, $f(x)$ is a convex function. Both of them depend on $R^{m\times n}$. $\beta(x) = f(x) - \mu \log\left(g(x)\right)$ where μ is a positive-valued constant, and condition is on the set $S = \{x: g(x) > 0\}$.
For $f(x)$: $f(\lambda\bar{x} + (1-\lambda)\hat{x}) \le \lambda f(\bar{x}) + (1-\lambda) f(\hat{x})$
For $g(x)$: $g(\lambda\bar{x} + (1-\lambda)\hat{x}) \ge \lambda g(\bar{x}) + (1-\lambda) g(\hat{x})$
Set $h(t) =- \mu \log(t)$, $h\left(g(x)\right) =- \mu \log\left(g(x)\right)$. $\frac{\partial h}{\partial x} =- \frac{\mu g^{'}(x)}{g(x)}$.

$\frac{\partial^2 h}{\partial^2 x} =- \frac{\mu g^{'}(x)}{g(x)} =- \frac{\mu(g^{''}(x)g(x) - \left(g^{'}(x)\right)^2)}{\left(g(x)\right)^2}$. Then we can analyze the result: since g(x) is concave, $g^{''}(x) \le 0 \rightarrow g^{''}(x)g(x) \le 0 \rightarrow \mu(g^{''}(x)g(x) - \left(g^{'}(x)\right)^2 \le 0 \rightarrow - \frac{\mu(g^{''}(x)g(x) - \left(g^{'}(x)\right)^2)}{\left(g(x)\right)^2}$. $\frac{\partial^2 \beta}{\partial^2 x} = f^{''}(x) + g^{''}(x)$, since $f(x)$ is a convex function, $f^{''}(x) \ge 0 \rightarrow f^{''}(x) + g^{''}(x) \ge 0 \rightarrow \frac{\partial^2 \beta}{\partial^2 x} \ge 0$. Thus, $\beta(x)$ is convex function in this case.

4. We know that $\left|\left|x\right|\right|_2^2 = x^T x$. Then proved by:

$$f(x) = \sum_{i,j=1}^{n} x_i x_j = x_k^2 + \sum_{j\neq 1}^{n} x_k x_j + \sum_{i\neq 1}^{n} x_i x_k$$

Hence, for

$$\frac{\partial h}{\partial k} = 2x_k + \sum_{j\neq 1}^{n} x_j + \sum_{i\neq 1}^{n} x_i$$

$$= 2x_k + 2\textstyle\sum_{j\neq 1}^{n} x_j \ \ (a_{ik} = a_{kj}, \text{here } a = I)$$

$$= 2\sum_{i,j=1}^{n} x_j = 2x$$

5. (a) We know that $f(x) = x_1^2 + x_2^2 + 2x_3^2 - x_1x_2 - x_2x_3 - x_1x_3, x \in R^3$. And we should transform to the form $f(x) = x^T A x$, For matrix A, the diagnose of A is the quadratic coefficient, and other is the half of the coefficient of the first degree.

So $A = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 2 \end{bmatrix}$, $f(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(b) $\nabla f(x)=\begin{bmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \frac{\partial f}{\partial x_3}\end{bmatrix}=\begin{bmatrix}2x_1-x_2-x_3\\ -x_1+2x_2-x_3\\ -x_1-x_2+4x_3\end{bmatrix}$

(c) $\nabla f(x)^2=\begin{bmatrix}\frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_1 x_3}\\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial^2 x_2} & \frac{\partial^2 f}{\partial x_2 x_3}\\ \frac{\partial^2 f}{\partial x_3 x_1} & \frac{\partial^2 f}{\partial x_3 x_2} & \frac{\partial^2 f}{\partial^2 x_3}\end{bmatrix}=\begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 4\end{bmatrix}$

(d) Calculate the determinant of $A-\lambda I. A-\lambda I=\begin{bmatrix}2-\lambda & -1 & -1\\ -1 & 2-\lambda & -1\\ -1 & -1 & 4-\lambda\end{bmatrix}$. Thus,

$\det(A-\lambda I)=\ (2-\lambda)(2-\lambda)(4-\lambda)-2-(2-\lambda)-(4-\lambda)-(2-\lambda)=0.$ Solve this equality and we can get $\lambda_1=3,\lambda_2=0.4384,\lambda_3=4.5616$. All of the eigenvalue are larger than zero, so $\nabla^2 f(x)$ is positive semi-definite matrix which means that $f(x)$ is convex function.

6. We know that $A\in R^{m\times n},\ x,b\in R,\ Q(x)=\left||Ax-b|\right|_2^2$

   (a) $Q(x)=(Ax-b)^T(Ax-b)=((Ax)^T b^T)(Ax-b)=(x^TA^T-b^T)(Ax-b)$
   $=\ x^TA^TAx-x^TA^Tb-b^TAx+b^Tb==x^TA^TAx-2Ab^Tx+b^Tb$

   Since $(x^TAx)^{'}=2Ax,\ \frac{\partial(x^TA^TAx)}{\partial x}=2A^TAx.$

   Since $(Ax)^{'}=A,\ \frac{\partial(2Ab^Tx)}{\partial x}=2Ab^T$

   Thus, $\nabla Q=2A^TAx-2Ab^T$.

   (b) When $\nabla f=0,\ x=(A^TA)^{-1}Ab^T$. Since $A\in R^{m\times n},\ A^TA$ is not a singular matrix, which means that it must exist solution in this case.

7. We know that $A=\begin{bmatrix}4 & \alpha\\ \alpha & 2\end{bmatrix}$, we should calculate the determinant of $A-\lambda I$.

   $A-\lambda I=\begin{bmatrix}4-\lambda & \alpha\\ \alpha & 2-\lambda\end{bmatrix}$, so $\det(A-\lambda I)=\ (4-\lambda)(2-\lambda)-\alpha^2=0$

   Simplify the equality, $\lambda^2-6\lambda+(8-\alpha^2)=0$. According to Square root formula of quadratic equation with one unknown, $\lambda=\frac{-b\pm\sqrt{b^2-4ac}}{2a}=\frac{6\pm\sqrt{36-(32-4\alpha^2)}}{2}$.

   Therefore, $\lambda_1=\frac{6+\sqrt{4+4\lambda}}{2},\ \lambda_2=\frac{6-\sqrt{4+4\lambda}}{2}$. If both of them are larger than 0, is the positive definite matrix and the positive semi-definite matrix. Since $\sqrt{4+4\lambda} \geq 0$, λ_1 must be larger than 0. For λ_2, if $\frac{6-\sqrt{4+4\lambda}}{2} = 0$, $\alpha = \pm 2\sqrt{2}$. If $\frac{6-\sqrt{4+4\lambda}}{2} > 0$, $\alpha \in (-2\sqrt{2}, 2\sqrt{2})$. To conclude, when $\alpha \in \left(-2\sqrt{2}, 2\sqrt{2}\right)$, A is the positive definite matrix and the positive semi-definite matrix. When $\alpha = \pm 2\sqrt{2}$, A is positive semi-definite matrix.