CompSci 206 Final

Xiaoying Li

12/08/2021

1.

(a) True

(b) True

(c) True

(d) False

(e) True

(f) True

(g) False

(h) True

2.

(a)

$$f(x_1, x_2) = x_1^2 - 4x_1 + 2x_2^2 + 7$$

$$\rightarrow \frac{\partial f}{\partial x_1} = 2x_1 - 4, \frac{\partial f}{\partial x_2} = 4x_2 \rightarrow \frac{\partial^2 f}{\partial x_1^2} = 2, \frac{\partial^2 f}{\partial x_2^2} = 4, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0, \frac{\partial^2 f}{\partial x_2 \partial x_1} = 0$$

$$\rightarrow \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

Let $H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$,

$$|H - \lambda I| = \begin{bmatrix} 2-\lambda & 0 \\ 0 & 4-\lambda \end{bmatrix} = (2-\lambda)(4-\lambda)$$

So, H’s eigenvalues are 2 and 4, which are both positive. Hence, H is positive semidefinite. Therefore, $f(x_1, x_2) = x_1 x_2$ is convex.

(b)

$$f(x,y)=\frac{x^2}{y}\rightarrow\frac{\partial f}{\partial x}=\frac{2x}{y},\frac{\partial f}{\partial y}=-\frac{x^2}{y^2}\rightarrow\frac{\partial^2 f}{\partial x^2}=\frac{2}{y},\frac{\partial^2 f}{\partial y^2}=\frac{2x^2}{y^3},\frac{\partial^2 f}{\partial x\partial y}=-\frac{2x}{y^2},\frac{\partial^2 f}{\partial y\partial x}=-\frac{2x}{y^2}$$

$$\rightarrow\nabla^2 f(x)=\begin{bmatrix}\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y}\\ \frac{\partial^2 f}{\partial y\partial x} & \frac{\partial^2 f}{\partial y^2}\end{bmatrix}=\begin{bmatrix}\frac{2}{y} & -\frac{2x}{y^2}\\ -\frac{2x}{y^2} & \frac{2x^2}{y^3}\end{bmatrix}=\frac{2}{y}\begin{bmatrix}1\\ -\frac{2x}{y}\end{bmatrix}\begin{bmatrix}1 & -\frac{2x}{y}\end{bmatrix}\succcurlyeq 0$$

Therefore, $f(x,y)=\frac{x^2}{y}\ with\ y>0$ is convex.

(c)

$f(\boldsymbol{x})=e^{-\|x\|^2}\ with\ \boldsymbol{x}\in R^n$ has positive semi-definite Hessian at some points (far from the origin) and negative semi-definite Hessian at some other points (close to the origin). Therefore, for $\boldsymbol{x}\in R^n$, $f(\boldsymbol{x})$ is neither convex nor concave.

(d)

$$f(\boldsymbol{x})=\|A\boldsymbol{x}-b\|_1$$

f is the composition of a norm, which is convex, and an affine function.

(e)

$$f: S_{++}^n\rightarrow R\ with\ f(X)=-\log\det(X)$$
$$\rightarrow -f: S_{++}^n\rightarrow R\ with\ f(X)=\log\det(X)$$

Define

$$g: R\rightarrow R, g(t)=f(X+tV)\ with\ dom\ g=\{t\mid X+tX\succ 0\}, for\ any\ X\succ 0\ and\ V\in S^n.$$

$$g(t)=\log\det(X+tX)=\log\det\left(X^{\frac{1}{2}}(I+tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X^{\frac{1}{2}}\right)=\sum_{i=1}^{n}\log(1+t\lambda_i)+\log\det X$$

Where $\lambda_1,\cdots,\lambda_n$ are the eigenvalues of $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$. Hence g is concave for any $X\succ 0$ and $V\in S^n$, so is -f.

Since -f is concave, therefore, f is convex.

3.

(a)

$$\text{minimize } f(\boldsymbol{x}) = -\sum_{i=1}^{n} \log x_i$$

$$\text{subject to } A\boldsymbol{x} = b$$

$$\text{with dom } f = R_{++}^{n} \text{ and } A \in R^{p\times n}, \text{and implicit contraint } x \succ 0$$

Using

$$f^*(y) = \sum_{i=1}^{n} (-1 - \log(-y_i)) = -n - \sum_{i=1}^{n} \log(-y_i)$$

$$\text{with dom } f^* = -R_{++}^{n}$$

The dual problem is

$$\text{maximize } g(v) = -b^T v + n + \sum_{i=1}^{n} \log\,(A^T v)_i$$

$$\text{with implicit contraint } A^T v \succ 0$$

(b)

From (a) we obtain, the dual feasibility equation can be solved by finding the x that minimizes

$$L(x, v)\colon \nabla f(x) + A^T v = -\left(\frac{1}{x_1}, \cdots \frac{1}{x_n}\right) + A^T v = 0$$

The KKT conditions are:

$$Ax = b, -\frac{1}{x} + A^T v = 0$$

(c)

The Newton's step for solving the primal problem with a feasible start x is given by solution V of the following linear equation:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

$$\text{where } \nabla f(x) = -\frac{1}{x}, \nabla^2 f(x) = diag(\frac{1}{x^2})$$

4.

(a)

$$f(\boldsymbol{x}) = f(x_1, x_2) = (x_1 + x_2^2)^2 = x_1^2 + 2x_1x_2^2 + x_2^4$$
$$\rightarrow \frac{\partial f}{\partial x_1} = 2x_1 + 2x_2^2, \frac{\partial f}{\partial x_2} = 4x_1x_2 + 4x_2^3$$
$$\rightarrow \nabla f(x) = \begin{bmatrix} 2x_1 + 2x_2^2 \\ 4x_1x_2 + 4x_2^3 \end{bmatrix} = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + 4x_2^2) \end{bmatrix}$$

(b)

$$\boldsymbol{x}^{(0)} = [0,1]^T, \nabla f\left(\boldsymbol{x}^{(0)}\right) = [2,4]^T$$
$$\rightarrow -\nabla f\left(\boldsymbol{x}^{(0)}\right) = [-2,-4]^T \ \textit{is the steepest decent direction}$$

Since

$$\boldsymbol{p}\left(-\nabla f\left(\boldsymbol{x}^{(0)}\right)\right) = [1,-1]^T \cdot [-2,-4]^T = 2 > 0$$

Therefore, $\boldsymbol{p}$ is a descent direction.

(c)

$$\textit{minimizes } f(\boldsymbol{x}^{(0)} + \alpha\boldsymbol{p})$$
$$\boldsymbol{x}^{(0)} + \alpha\boldsymbol{p} = [0,1]^T + \alpha[-1,1]^T = [\alpha, 1-\alpha]^T$$
$$\rightarrow f\left(\boldsymbol{x}^{(0)} + \alpha\boldsymbol{p}\right) = (\alpha^2 + (1-\alpha)^2)^2 = \alpha^2 + (1-\alpha)^4 + 2\alpha(1-\alpha)^2$$
$$\rightarrow \frac{\partial f\left(\boldsymbol{x}^{(0)} + \alpha\boldsymbol{p}\right)}{\partial \alpha} = 4\alpha^3 - 6\alpha^2 + 6\alpha - 2 = 0 \rightarrow \alpha = \frac{1}{2}$$
$$\rightarrow f\left(\boldsymbol{x}^{(0)} + \alpha\boldsymbol{p}\right) = (\frac{1}{2})^2 + (1-\frac{1}{2})^4 + 2\cdot\frac{1}{2}(1-\frac{1}{2})^2 = \frac{9}{16}$$

(d)

$$\frac{\partial f}{\partial x_1} = 2x_1 + 2x_2^2, \frac{\partial f}{\partial x_2} = 4x_1x_2 + 4x_2^3$$
$$\rightarrow \frac{\partial^2 f}{\partial x_1^2} = 2, \frac{\partial^2 f}{\partial x_2^2} = 4x_2 + 12x_2^2, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 4x_2, \frac{\partial^2 f}{\partial x_2 \partial x_1} = 4x_2$$

$$\rightarrow \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 4x_2 \\ 4x_2 & 4x_2 + 12x_2^2 \end{bmatrix}$$

(e)

$$\Delta \boldsymbol{x}^{(0)} = -\nabla^2 f(x)^{-1} \nabla f(x) = -\begin{bmatrix} 2 & 4 \\ 4 & 12 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\rightarrow \boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + t\Delta \boldsymbol{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\rightarrow f\left(\boldsymbol{x}^{(1)}\right) = (-1 + 1^2) = 0$$

5.

(a)

$$min_{\boldsymbol{x} \in R^n} \|\boldsymbol{x}\|_2^2, subject\ to \sum_{i=1}^{n} x_i = 1 \quad (1)$$

The objective is a convex quadratic function, and all constrains are affine. The Lagrangian of this problem is given by

$$\mathcal{L}(\mathbf{x}, \mu) = \boldsymbol{x}^T \boldsymbol{x} + \mu(\mathbf{1}^T \boldsymbol{x} - 1)$$

with optimality conditions found by taking partial derivatives with respect to **x**.

$$\nabla_X \mathcal{L}(\mathbf{x}, \mu) = 2\boldsymbol{x} + \mathbf{1}\mu = 0$$

$$\rightarrow \boldsymbol{x} = -\frac{1}{2}\mathbf{1}\mu \quad (2)$$

By combining (2) with the constraint in (1)

$$1 = \mathbf{1}^T X = -\frac{1}{2}\mathbf{1}^T \mathbf{1}\mu$$

such that $\mu^* = -\frac{2}{n}$ and $\boldsymbol{x}^* = \frac{1}{n}\mathbf{1}$

(b)

$$\min_{x \in R^n} \boldsymbol{x}^T Q_1 \boldsymbol{x}\, , subject\ to\ \boldsymbol{x}^T \boldsymbol{Q}_1 \boldsymbol{x} = 1$$

Before solving this problem, we will first introduce an equivalent optimization problem

$$\min_{x \in R^n} \boldsymbol{x}^T \boldsymbol{Q}_1 \boldsymbol{x}\,, subject\ to\ \boldsymbol{y}^T \boldsymbol{y} = 1, \boldsymbol{y} = \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{x} \quad (3)$$

where we exploit the fact that $\boldsymbol{Q}_2$ is positive definite and choose $\boldsymbol{Q}_2^{\frac{1}{2}}$ such that $\boldsymbol{Q}_2^{\frac{1}{2}} = \boldsymbol{Q}_2^{T\frac{1}{2}}$. The Lagrangian of this new problem is given by

$$\mathcal{L}(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \mathrm{v}) = \boldsymbol{x}^T \boldsymbol{Q}_1 \boldsymbol{x} + v(\boldsymbol{y}^T \boldsymbol{y} - 1) + \boldsymbol{\mu}^T (\boldsymbol{y} - \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{X})$$

With optimality condition found by taking partial derivative with respect to X and Y.

$$\nabla_X \mathcal{L}(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \mathrm{v}) = 2\boldsymbol{Q}_1 \boldsymbol{x} - \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\mu} = 0$$

$$\nabla_Y \mathcal{L}(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \mathrm{v}) = \boldsymbol{\mu} + 2v\boldsymbol{y} = 0$$

$$\rightarrow \boldsymbol{x} = \frac{1}{2} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\mu}$$

$$Y = -\frac{1}{2v} \boldsymbol{\mu}$$

Substituting (4) into the constraints in (3), we obtain,

$$\boldsymbol{y} = \frac{1}{2} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\mu} = -\frac{1}{2v} \boldsymbol{\mu} \quad (4)$$

Thus $-\frac{1}{2v}$ and $\boldsymbol{\mu}$ form and eigenpair of the matrix $\boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\mu}$. By then substituting (4) into the objective of (3) it can be shown that

$$\boldsymbol{x}^T \boldsymbol{Q}_1 \boldsymbol{x} = \frac{1}{4} \boldsymbol{\mu}^T \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\mu}$$

Such that the objective is minimized when $\boldsymbol{\mu}$ is the eigenvector of $\boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}}$ corresponding to the smallest eigenvalue. We will denote this vector by $\boldsymbol{\mu}^* = \boldsymbol{\alpha}_{min}$. From the final constraint $\boldsymbol{y}^T \boldsymbol{y} = \frac{\boldsymbol{\mu}^T \boldsymbol{\mu}}{4v^2} = 1$ we can determine the value of $v^* = \mp 2$.

Finally, $\boldsymbol{x}^* = \mp \frac{1}{2} \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2^{\frac{1}{2}} \boldsymbol{\alpha}_{min}$ and $\boldsymbol{y}^* = \mp \boldsymbol{\alpha}_{min}$.

6.

(a)

Since $\mathbf{1}$ is the vector of all ones, $\boldsymbol{x} \in R^n$ is also a vector, we obtain

$$\|\boldsymbol{x}\|_\infty = \max(|\boldsymbol{x}_i|) \leq 1 \rightarrow all\ |\boldsymbol{x}_i| \leq 1 \rightarrow \mathbf{1} \leq \boldsymbol{x} \leq \mathbf{1}$$

Therefore, problem P is equivalent to problem $\hat{P}$.

(b)

The Langrangian function for $\hat{P}$ is

$$\mathcal{L}(\mathbf{x}, \lambda) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T\mathbf{x} + \lambda_1(\boldsymbol{x} - \mathbf{1}) + \lambda_2(-\boldsymbol{x} - \mathbf{1}), with\ \lambda_1 \geq 0\ and\ \lambda_2 \geq 0$$

(c)

$$\frac{\partial \mathcal{L}}{\partial x} = Q\boldsymbol{x} + \boldsymbol{c} + \lambda_1\mathbf{1} - \lambda_2\mathbf{1} = 0$$

$$\boldsymbol{x}^* = Q^{-1}(-\lambda_1\mathbf{1} + \lambda_2\mathbf{1} - \boldsymbol{c}), define\ \boldsymbol{y} = -\lambda_1\mathbf{1} + \lambda_2\mathbf{1}$$

Substitution into $\mathcal{L}$, we obtain,

$$\mathcal{L}(\boldsymbol{x}^*, \mathbf{y}, \lambda) = \frac{1}{2}[Q^{-1}(\mathbf{y} - \mathbf{c})]^T Q[Q^{-1}(\mathbf{y} - \mathbf{c})] + \boldsymbol{c}^T Q^{-1}(\mathbf{y} - \mathbf{c}) - (\lambda_1 - \lambda_2)\boldsymbol{x} - (\lambda_1 + \lambda_2)\mathbf{1}$$

$$= \frac{1}{2}[Q^{-1}(\mathbf{y} - \mathbf{c})]^T Q[Q^{-1}(\mathbf{y} - \mathbf{c})] + \boldsymbol{c}^T Q^{-1}(\mathbf{y} - \mathbf{c}) - \mathbf{y}^T Q^{-1}(\mathbf{y} - \mathbf{c}) - \|\mathbf{y}\|_1$$

$$\mathbf{y} = -\lambda_1\mathbf{1} + \lambda_2\mathbf{1} = [-\lambda_1, \lambda_2]^T\mathbf{1}$$

$$\rightarrow g(\mathbf{y}) = \frac{1}{2}(\mathbf{y} - \mathbf{c})^T Q^{-1}(\mathbf{y} - \mathbf{c}) + \boldsymbol{c}^T Q^{-1}\mathbf{y} - \boldsymbol{c}^T Q^{-1}\mathbf{c} - \boldsymbol{y}^T Q^{-1}\mathbf{y} + \boldsymbol{y}^T Q^{-1}\mathbf{c} - \|\mathbf{y}\|_1$$

$$= \frac{1}{2}(\mathbf{y} - \mathbf{c})^T Q^{-1}(\mathbf{y} - \mathbf{c}) - (\mathbf{y} - \mathbf{c})^T Q^{-1}(\mathbf{y} - \mathbf{c}) - \|\mathbf{y}\|_1$$

$$= -\frac{1}{2}(\mathbf{y} - \mathbf{c})^T Q^{-1}(\mathbf{y} - \mathbf{c}) - \|\mathbf{y}\|_1$$

Hence, the Lagrangian dual for $\hat{P}$ is the problem D, and this is also the Langrangian dual for P.