/

APM462 HW1 Rui Qiu #999292509

Exercise 1

a. Solution $f(a)=\int_0^1 (g(x)-p_a(x))^2 dx$

Let $w=(1, x, x^2, \cdots, x^n)$, $a=(a_0, a_1, \dots, a_n)$ both are column vectors, $f(a)=\int_0^1 (g(x)-a^Tw)^2 dx$

$$=\int_0^1 (g(x)^2-2a^Twg(x)+(a^Tw)^2)dx$$

$$=\int_0^1 (g(x))^2 dx - 2a^T\int_0^1 w\,g(x)dx+\int_0^1 a^Tww^Ta\,dx$$

$$=a^T\left(\int_0^1 ww^T dx\right)a-2a^T\int_0^1 wg(x)dx+\int_0^1 (g(x))^2 dx$$

$$=a^TQa-2b^Ta+\int_0^1 (g(x))^2 dx$$

Therefore $Q=\int_0^1 w.w^T dx$, $b=\int_0^1 wg(x)dx$, $c=\int_0^1 g(x)^2 dx$

Thus $f(a)=a^TQa-2b^Ta+c$

Notice that $Q=\int_0^1 w\cdot w^T dx$

Since $w\cdot w^T=\begin{bmatrix}1\\x\\x^2\\\vdots\\x^n\end{bmatrix}[1\ x\ x^2\cdots x^n]=\begin{bmatrix}1 & x & x^2 & \cdots & x^n\\ x & x^2 & x^3 & & \vdots\\ x^2 & & & & \vdots\\ \vdots & & & & \\ x^n & & \cdots & & x^{2n}\end{bmatrix}$

So $Q=\begin{bmatrix}1 & \frac{1}{2} & \cdots & \frac{1}{n+1}\\ \frac{1}{2} & \frac{1}{3} & & \vdots\\ \vdots & & & \vdots \\ \frac{1}{n+1} & \cdots & & \frac{1}{2n+1}\end{bmatrix}$

Then Q is symmetric.

b. Solution By a we know $f(a) = a^TQa - 2b^Ta + C$

First order necessary condition:

$\nabla f(a) = 0 = 2a^TQ - 2b^T$

If a^* is minimum point, then

$2a^{*T}Q = 2b^T$

so $Q \cdot a^* = b$

c. Since $g(x) \equiv 0$, then by (a). we know $b = 0$, $C = 0$

$f(a) = a^TQa$.

so by (b). $Qa^* = b = 0$

$$\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n+1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n+1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

So $a_0 = a_1 = \cdots = a_n = 0$ is a solution.

And we want to prove it's the only solution.

we have $(n+1)$ unknown variables.

and $(n+1)$ independent equations:

$$\begin{cases} a_0 + \frac{1}{2}a_1 + \cdots + \frac{1}{n+1}a_n = 0 \\ \frac{1}{2}a_0 + \frac{1}{3}a_1 + \cdots + \frac{1}{n+2}a_n = 0 \\ \vdots \\ \frac{1}{n+1}a_0 + \frac{1}{n+2}a_1 + \cdots + \frac{1}{n+3}a_n = 0 \end{cases}$$

Thus they must have one and only one solution.

Note: e.g. $\begin{cases} x+2y=0 \\ 2x+4y=0 \end{cases}$ does not have only one solution,

but $\begin{cases} x+2y=0 \\ x-y=0 \end{cases}$ has only one solution.

d. Solution:

$\nabla f(a) = 2Qa - 2b^T$

$\nabla^2 f(a) = 2Q$

By 2nd order Taylor expansion in E^n.

$$f(x+v) = f(x) + \nabla f(x) \cdot v + \frac{1}{2} v^T \nabla f(x) v + o(|v|^2)$$

Let a^* satisfy $\nabla f(a^*) = 0$.

then $f(a^*+v) = f(a^*) + \nabla f(a^*) \cdot v + \frac{1}{2} v^T \nabla f(a^*) v + o(|v|^2)$

$$= f(a^*) + 0 \cdot v + v^T Q v + o(|v|^2)$$

Suppose $v^T Q v$ negative definite, ie. $v^T Q v < 0$.

Then $f(a^*+v) < f(a^*)$ for some v.

Contradiction, since a^* is a minimum point.

So $v^T Q v \geq 0$, nonnegative definite. (positive semidefinite)

Then if $v^T Q v = 0$

$\Rightarrow$ Q has at least one eigenvalue say $\lambda = 0$.

then $\det Q = 0$

Impossible since Q is invertible.

Hence $v^T Q v > 0$.

i.e. Q is positive definite.

Exercise 2.

Solution $\nabla f(x)=0$

$$\frac{\partial f}{\partial x}=4x+y-6=0$$

$$\frac{\partial f}{\partial x}=x+2y+z-8=0$$

$$\frac{\partial f}{\partial z}=y+2z-8=0$$

So $x=1, y=2, z=3$

$\nabla f(1,2,3)=0$

$$F(1,2,3)=\nabla^2 f(1,2,3)=\begin{bmatrix}4 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 2\end{bmatrix}$$

For any feasible direction $d=[a, b\ c]$

$$d^T\nabla^2 f(x)d=[a\ b\ c]\begin{bmatrix}4 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 2\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}$$

$$=4a^2+2ab+2b^2+2bc+2c^2$$

$$=3a^2+(a+b)^2+(b+c)^2+c^2\geq 0$$

Hence $(1,2,3)$ is a local minimum.

Next we want to know it's global.

Need find eigenvalues of $F(1,2,3)$

$$F-\lambda I=\begin{bmatrix}4-\lambda & 1 & 0\\ 1 & 2-\lambda & 1\\ 0 & 1 & 2-\lambda\end{bmatrix}=(4-\lambda)[(2-\lambda)^2-1]-(2-\lambda)+0$$

$$=(4-\lambda)[4-4\lambda+\lambda^2-1]-(2-\lambda)$$

$$=12-16\lambda+4\lambda^2-3\lambda+4\lambda^2-\lambda^3-2+\lambda$$

$$=-\lambda^3+8\lambda^2-18\lambda+10$$

It's continuous on $\mathbb{R}$. (3 roots possibly) let it be $g(\lambda)$

We notice that $g(0)>0$, $g(1)=-1<0$.

By intermedia value theorem, $\lambda_1\in(0,1)$

Similarly we can find $\lambda_2\in(2,3)$, $\lambda_3\in(4,5)$

For λ_i, $i\in\{1,2,3\}$, $g(\lambda_i)=0$

Since $\lambda_1, \lambda_2, \lambda_3$ all positive then $F(x,y,z)$ is positive definite.
By theorem, $(1,2,3)$ is a global minimum.

Exercise 3.

a. Solution: Since g is a convex function.

then $g(\theta x+(1-\theta)y) \leq \theta(g(x)) + (1-\theta)g(y)$.

Since LHS $\leq$ RHS, f is a non-decreasing function,

$f(LHS) \leq f(RHS)$

i.e. $f(\theta x+(1-\theta)y) \leq f(\theta(g(x)) + (1-\theta)g(y))$

$$= f(\theta(g(x)) + f((1-\theta)g(y))$$

$$= \theta f(g(x)) + (1-\theta)f(g(y))$$

$$F(\theta x+(1-\theta)y) \leq \theta F(x) + (1-\theta)F(y)$$

b. Solution: Idea: want to show $H = \nabla^2 F$ is nonnegative-definite

need to show $\forall d$, $d^T H d \geq 0$

Suppose $d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \in E^n$ (H is $n \times n$)

$$H = \begin{bmatrix} \frac{\partial^2 F}{\partial^2 x_1} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial^2 x_2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 F}{\partial^2 x_n} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & & & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix}$$ for simplicity

(see next page)

$$d^T H d = (d_1^2 h_{11} + \cdots + d_1 d_n h_{n1}) + (d_1 d_2 h_{12} + \cdots + d_n d_2 h_{n2})$$
$$+ \cdots + (d_1 d_n h_{1n} + \cdots + d_n^2 h_{nn})$$
$$= \sum_{i,j=1}^{n} d_i d_j h_{ij} = \Sigma_{i,j=1}^{n} d_i d_j \left(f'' \frac{\partial g}{\partial x_i} \cdot \frac{\partial g}{\partial x_j} + f' \cdot \frac{\partial^2 g}{\partial x_i \partial x_j} \right)$$

Since f is nondecreasing. then $f' \geq 0$
Since g is convex, then $\nabla^2 g \geq 0$

so $d^T \nabla^2 g d \geq 0$
i.e. $\Sigma_{i,j=1}^{n} \frac{\partial^2 g}{\partial x_i \partial x_j} d_i d_j \geq 0$

As $f' \geq 0$

Then $\Sigma_{i,j=1}^{n} \left(f' \frac{\partial^2 g}{\partial x_i \partial x_j} \right) d_i d_j \geq 0$ ①

Since f is also convex, $\nabla^2 f \geq 0$

so $f'' \geq 0 \Rightarrow \Sigma_{i,j=1}^{n} d_i d_j \left(f'' \frac{\partial g}{\partial x_i} \frac{\partial g}{\partial x_j} \right) \geq 0$ ②

Next, we want to show $\Sigma_{i,j=1}^{n} \frac{\partial g}{\partial x_i} \frac{\partial g}{\partial x_j} d_i d_j \geq 0$

$\Rightarrow \Sigma_{i,j=1}^{n} \frac{\partial g}{\partial x_i} d_i \cdot \Sigma_{i,j=1}^{n} \frac{\partial g}{\partial x_j} d_j$

$= (\nabla g d^T) \cdot (\nabla g d^T)$

$= (\nabla g d^T)^2 \geq 0$ ③

Therefore, consider ① ② & ③, we can say $d^T H d \geq 0$
i.e. H is positive semidefinite.

Exercise 4.

Solution:

$\forall\ 0 \le \alpha \le 1,\ \forall x_1, x_2 \in E^n$

① $g(\alpha x_1 + (1-\alpha)x_2)$

$= \max\{f_1(\alpha x_1 + (1-\alpha)x_2), \cdots, f_k(\alpha x_1 + (1-\alpha)x_2)\}$

$\le \max\{\alpha f_1(x_1) + (1-\alpha)f_1(x_2), \cdots, \alpha f_k(x_1) + (1-\alpha)f_k(x_2)\}$

② $\alpha g(x_1) + (1-\alpha)g(x_2)$

$= \alpha \max\{f_1(x_1), \cdots, f_k(x_1)\} + (1-\alpha)\max\{f_1(x_2), \cdots, f_k(x_2)\}$

$= \max\{\alpha f_1(x_1), \cdots, \alpha f_k(x_1)\} + \max\{(1-\alpha)f_1(x_2), \cdots, (1-\alpha)f_k(x_2)\}$

$= \max\{\alpha f_1(x_1) + (1-\alpha)f_1(x_2), \cdots, \alpha f_k(x_1) + (1-\alpha)f_k(x_2)\}$

Hence ① $\le$ ②

Thus $g(x)$ is convex.