3장 연습문제 해답

01.

n에 대한 수학적귀납법을 이용해 증명해봅시다.

① <u>$n=2$일 때</u>

볼록집합의 정의에 의해, $n=2$일때는 당연히 성립합니다.

② <u>$n=k$일 때 성립한다고 가정</u>

주어진 명제가 $n=k$일 때 성립한다고 가정해봅시다. 즉, $\theta_i \geq 0$이고 $\theta_1+\cdots+\theta_k=1$인 실수 $\theta_1,\cdots,\theta_k$에 대하여 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k \in C$이면 $\theta_1\boldsymbol{x}_1+\cdots+\theta_k\boldsymbol{x}_k \in C$입니다. 이제, $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k,\boldsymbol{x}_{k+1} \in C$ 일 때, $\theta_i \geq 0$이고 $\theta_1+\cdots+\theta_k+\theta_{k+1}=1$을 만족하는 실수에 대해 $\theta_1\boldsymbol{x}_1+\cdots+\theta_k\boldsymbol{x}_k+\theta_{k+1}\boldsymbol{x}_{k+1}$을 생각해보면 다음과 같이 쓸 수 있습니다.

$$\theta_1\boldsymbol{x}_1+\cdots+\theta_k\boldsymbol{x}_k+\theta_{k+1}\boldsymbol{x}_{k+1}=(1-\theta_{k+1})\left(\frac{\theta_1}{1-\theta_{k+1}}\boldsymbol{x}_1+\cdots+\frac{\theta_k}{1-\theta_{k+1}}\boldsymbol{x}_k\right)+\theta_{k+1}\boldsymbol{x}_{k+1}$$

여기서, $\dfrac{\theta_1}{1-\theta_{k+1}}+\cdots+\dfrac{\theta_k}{1-\theta_{k+1}}=\dfrac{\theta_1+\cdots+\theta_k}{1-\theta_{k+1}}=\dfrac{1-\theta_{k+1}}{1-\theta_{k+1}}=1$이므로 가정에 의해

$$\frac{\theta_1}{1-\theta_{k+1}}\boldsymbol{x}_1+\cdots+\frac{\theta_k}{1-\theta_{k+1}}\boldsymbol{x}_k \in C$$

입이고, C는 볼록집합이므로 볼록집합의 정의에 의해 다음과 같습니다.

$$\theta_1\boldsymbol{x}_1+\cdots+\theta_k\boldsymbol{x}_k+\theta_{k+1}\boldsymbol{x}_{k+1}=(1-\theta_{k+1})\left(\frac{\theta_1}{1-\theta_{k+1}}\boldsymbol{x}_1+\cdots+\frac{\theta_k}{1-\theta_{k+1}}\boldsymbol{x}_k\right)+\theta_{k+1}\boldsymbol{x}_{k+1} \in C$$

따라서 수학적 귀납법에 의해 주어진 명제는 참입니다.

02.

(a) $\|\boldsymbol{x}_1\|_2 \leq t_1$, $\|\boldsymbol{x}_2\|_2 \leq t_2$를 만족하는 집합 C의 임의의 두 원소 $(\boldsymbol{x}_1,t_1),(\boldsymbol{x}_2,t_2)$와 $0 \leq \theta \leq 1$인 임의의 θ에 대해,

$$\begin{aligned}\|\theta\boldsymbol{x}_1+(1-\theta)\boldsymbol{x}_2\|_2 &\leq |\theta|\,\|\boldsymbol{x}_1\|_2+|1-\theta|\,\|\boldsymbol{x}_2\|_2\\ &\leq \theta\|\boldsymbol{x}_1\|_2+(1-\theta)\|\boldsymbol{x}_2\|_2\\ &\leq \theta t_1+(1-\theta)t_2\end{aligned}$$

이 성립합니다. 따라서 $\theta(\boldsymbol{x}_1, t_1) + (1-\theta)(\boldsymbol{x}_2, t_2) = (\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \theta t_1 + (1-\theta)t_2) \leq C$이므로 집합 C는 볼록집합입니다.

(b) 주어진 집합이 볼록집합일 필요충분조건은 임의의 직선 $\{\boldsymbol{x} + t\boldsymbol{v} | t \in \mathbb{R}\}$과의 주어진 집합과의 교집합이 볼록집합인 것입니다. (직접 증명해봅시다!) 주어진 집합 C와 임의의 직선 $\{\boldsymbol{x} + t\boldsymbol{v} | t \in \mathbb{R}\}$과의 교점을 생각하면,

$$(\boldsymbol{x} + t\boldsymbol{v})^T A(\boldsymbol{x} + t\boldsymbol{v}) + \boldsymbol{b}^T(\boldsymbol{x} + t\boldsymbol{v}) + c = (\boldsymbol{v}^T A\boldsymbol{v})t^2 + (\boldsymbol{b}^T\boldsymbol{v} + 2\boldsymbol{x}^T A\boldsymbol{v})t + c + \boldsymbol{b}^T\boldsymbol{x} + \boldsymbol{x}^T A\boldsymbol{x} \leq 0$$

을 만족하는 $\boldsymbol{x} + t\boldsymbol{v}$의 집합이고, 이 집합이 볼록집합이기 위해서는 t^2의 계수인 $\boldsymbol{v}^T A\boldsymbol{v}$가 음이 아닌 실수, 즉 $\boldsymbol{v}^T \boldsymbol{A}\boldsymbol{v}$여야 합니다. 한편, 주어진 문제에서 A는 준정부호 행렬이므로 임의의 $\boldsymbol{v}$에 대해 $\boldsymbol{v}^T A\boldsymbol{v} \geq 0$이고, 따라서 주어진 집합 C는 볼록집합입니다.

03.

(a) 집합 $C_1 \cap C_2$의 임의의 두 원소 x_1, x_2는 집합 C_1과 C_2 모두 속하는 원소이고, C_1, C_2가 볼록집합이므로 볼록집합의 정의에 의해 $0 \leq \theta \leq 1$인 실수 θ에 대해

$$\theta x_1 + (1-\theta)x_2 \in C_1, \ \theta x_1 + (1-\theta)x_2 \in C_2$$

따라서 $\theta x_1 + (1-\theta)x_2 \in C_1 \cap C_2$이므로 집합 $C_1 \cap C_2$는 볼록집합입니다.

(b) 볼록집합 C와 아핀함수 $f(\boldsymbol{x}) = A\boldsymbol{x} + b$에 대해 정의된 집합 $f(C)$에 속하는 임의의 원소 $\boldsymbol{y}_1$, $\boldsymbol{y}_2$는 어떤 $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \mathbb{R}^n$에 대해 다음이 성립합니다.

$$\boldsymbol{y}_1 = f(\boldsymbol{x}_1) = A\boldsymbol{x}_1 + b, \ \boldsymbol{y}_2 = f(\boldsymbol{x}_2) = A\boldsymbol{x}_2 + b$$

$0 \leq \theta \leq 1$인 실수 θ와 위 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$에 대해, $f(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2)$를 생각해보면

$$\begin{aligned} f(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) &= A(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) + b \\ &= \theta A\boldsymbol{x}_1 + (1-\theta)A\boldsymbol{x}_2 + (\theta + 1 - \theta)b \\ &= \theta(A\boldsymbol{x}_1 + b) + (1-\theta)(A\boldsymbol{x}_2 + b) \\ &= \theta\boldsymbol{y}_1 + (1-\theta)\boldsymbol{y}_2 \end{aligned}$$

이므로, $\theta\boldsymbol{y}_1 + (1-\theta)\boldsymbol{y}_2 \in f(C)$입니다. 따라서, 집합 $f(C)$는 볼록집합입니다.

04.

❶ <u>f가 볼록함수 $\Rightarrow$ epi$f$가 볼록집합</u>

f가 볼록함수라고 하고, 집합 epi$f$의 임의의 두 원소 $(\boldsymbol{x}_1, t_1)$, $(\boldsymbol{x}_2, t_2)$를 생각합시다. 그러면 $0 \leq \theta \leq 1$인 임의의 θ에 대해 다음이 성립합니다.

$$\theta t_1 + (1-\theta)t_2 \geq \theta f(\boldsymbol{x}_1) + (1-\theta)f(\boldsymbol{x}_2) \geq f(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2)$$

따라서 $(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \theta t_1 + (1-\theta)t_2) = \theta(\boldsymbol{x}_1, t_1) + (1-\theta)(\boldsymbol{x}_2, t_2)$도 집합 epi$f$의 원소이므로, 집합 epi$f$는 볼록집합입니다.

❷ <u>epi$f$가 볼록집합 $\Rightarrow$ f가 볼록함수</u>

집합 epi$f$은 볼록집합이므로 집합 epi$f$의 $t_1 = f(\boldsymbol{x}_1)$, $t_2 = f(\boldsymbol{x}_2)$를 만족하는 두 원소 $(\boldsymbol{x}_1, t_1)$, $(\boldsymbol{x}_2, t_2)$와 $0 \leq \theta \leq 1$인 임의의 θ에 대해

$$\theta(\boldsymbol{x}_1, t_1) + (1-\theta)(\boldsymbol{x}_2, t_2) = (\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \theta t_1 + (1-\theta)t_2)$$

도 집합 epi$f$의 원소입니다. 즉, epi$f$의 정의에 의해

$$f(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \leq \theta t_1 + (1-\theta)t_2 = \theta f(\boldsymbol{x}_1) + (1-\theta)f(\boldsymbol{x}_2)$$

이므로, 함수 f는 볼록함수입니다.

05.

집합 X의 임의의 두 원소 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$는 볼록함수 $f_i(i=1,\cdots,m)$와 벡터 $\boldsymbol{a}_i$와 스칼라 b_i $(i=1,\cdots,p)$에 대해 다음이 성립합니다.

$$f_i(\boldsymbol{x}_1) \leq 0,\ \boldsymbol{a}_i^T \boldsymbol{x}_1 = b_i$$
$$f_i(\boldsymbol{x}_2) \leq 0,\ \boldsymbol{a}_i^T \boldsymbol{x}_2 = b_i$$

한편, $0 \leq \theta \leq 1$인 실수 θ에 대해, 볼록함수 f_i의 정의와 θ, $1-\theta \geq 0$는 사실을 이용하면 $f_i(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2)$와 $\boldsymbol{a}_i^T(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2)$는

$$f_i(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \leq \theta f(\boldsymbol{x}_1) + (1-\theta)f(\boldsymbol{x}_2) \leq 0$$
$$\boldsymbol{a}_i^T(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) = \theta \boldsymbol{a}_i^T \boldsymbol{x}_1 + (1-\theta)\boldsymbol{a}_i^T \boldsymbol{x}_2 = \theta b + (1-\theta)b = b$$

이므로 $\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in X$입니다. 따라서, 주어진 집합 X는 볼록집합입니다.

06.

(a) 볼록함수 f_i $(i=1,\cdots,n)$은 정의에 의해 $\mathrm{dom} f_i$에 속하는 임의의 원소 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$와 $0 \le \theta \le 1$인 실수 θ에 대해 $f_i(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \le \theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2)$가 성립합니다. 주어진 함수 $f = \sum_{i=1}^{n} w_i f_i$의 정의역은 $\mathrm{dom} f = \bigcap_{i=1}^{n} \mathrm{dom} f_i$이므로, $\mathrm{dom} f$에 속하는 임의의 원소 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$와 $0 \le \theta \le 1$인 실수 θ에 대해 다음이 성립합니다.

$$
\begin{aligned}
f(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) &= \sum_{i=1}^{n} w_i f_i(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \\
(\because w_i \ge 0) &\le \sum_{i=1}^{n} w_i (\theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2)) \\
&= \theta \sum_{i=1}^{n} w_i f_i(\boldsymbol{x}_1) + (1-\theta) \sum_{i=1}^{n} w_i f_i(\boldsymbol{x}_2) \\
&= \theta f(\boldsymbol{x}_1) + (1-\theta) f(\boldsymbol{x}_2)
\end{aligned}
$$

따라서, 함수 f는 볼록함수입니다.

(b) $\mathbb{R}^m$에 속하는 임의의 두 원소 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$와 $0 \le \theta \le 1$인 실수 θ에 대해,

$$
\begin{aligned}
g(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) &= f(A(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) + b) \\
&= f(\theta A\boldsymbol{x}_1 + (1-\theta) A\boldsymbol{x}_2 + (\theta + 1 - \theta) b) \\
&= f(\theta(A\boldsymbol{x}_1 + b) + (1-\theta)(A\boldsymbol{x}_2 + b))
\end{aligned}
$$

한편, f가 볼록함수이므로

$$
f(\theta(A\boldsymbol{x}_1 + b) + (1-\theta)(A\boldsymbol{x}_2 + b)) \le \theta f(A\boldsymbol{x}_1 + b) + (1-\theta) f(A\boldsymbol{x}_2 + b)
$$

가 성립하고,

$$
\begin{aligned}
g(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) &= f(\theta(A\boldsymbol{x}_1 + b) + (1-\theta)(A\boldsymbol{x}_2 + b)) \\
&\le \theta f(A\boldsymbol{x}_1 + b) + (1-\theta) f(A\boldsymbol{x}_2 + b) \\
&= \theta g(\boldsymbol{x}_1) + (1-\theta) g(\boldsymbol{x}_2)
\end{aligned}
$$

입니다. 따라서 함수 g는 볼록함수입니다.

07.

① **일계조건**

ⓐ (a) $\Rightarrow$ (b)

[정리 3-1]의 증명에서 $g(\theta)=f(\boldsymbol{x}+\theta(\boldsymbol{y}-\boldsymbol{x}))$로 두면,

$$\lim_{\theta\to 0}\frac{f(\boldsymbol{x}+\theta(\boldsymbol{y}-\boldsymbol{x}))-f(\boldsymbol{x})}{\theta}=\lim_{\theta\to 0}\frac{g(\theta)-g(0)}{\theta}=g'(0)$$

입니다. 연쇄법칙에 의해 $g'(\theta)=\nabla f(\boldsymbol{x}+\theta(\boldsymbol{y}-\boldsymbol{x}))^T(\boldsymbol{y}-\boldsymbol{x})$이므로, $g'(0)=\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})$이고, [정리 3-1]의 증명에서 이를 $f'(x)(y-x)$ 대신 대입하면 $n>1$인 경우가 증명됩니다.

ⓑ (b) $\Rightarrow$ (a)

[정리 3-1]의 증명에서 f'을 모두 ∇f로 바꾸면 $n>1$인 경우가 증명됩니다.

② **이계조건**

$n>1$일 때, 주어진 함수 f가 볼록함수인 조건과 동치조건은 임의의 $\boldsymbol{x}_0\in\mathrm{dom}f$와 $\boldsymbol{v}$에 대해서 $g(t)=f(\boldsymbol{x}_0+t\boldsymbol{v})$가 t에 대한 볼록함수인 것입니다. (직접 증명해봅시다!) 즉,

$$g''(t)=\boldsymbol{v}^T\nabla^2 f(\boldsymbol{x}_0+t\boldsymbol{v})\boldsymbol{v}$$

이 임의의 t와 $\boldsymbol{x}_0\in\mathrm{dom}f$, $\boldsymbol{v}$에 대해 음이 아닌 실수여야합니다. 이는 $\nabla^2 f(\boldsymbol{x})$가 양의 중정부호 행렬이라는 뜻이므로 $n>1$인 경우의 [정리 3-2]가 증명됩니다.

08.

(a) 함수 $g(\boldsymbol{x})=\|\boldsymbol{x}\|$로 정의하면, 주어진 함수 $f(\boldsymbol{x})=\|A\boldsymbol{x}-b\|$는 $f(\boldsymbol{x})=g(A\boldsymbol{x}-b)$로 볼 수 있습니다. 따라서 [연습문제 6-(b)]에 의해 주어진 함수 f는 볼록함수입니다.

(b) 함수 $h(\boldsymbol{x})=|x_1|^p+\cdots+|x_n|^p$으로 두면, 주어진 함수 $f(\boldsymbol{x},t)$는 다음과 같이 표현할 수 있습니다.

$$f(\boldsymbol{x},t)=th(\boldsymbol{x}/t)$$

따라서, $f(\boldsymbol{x},t)=th(\boldsymbol{x}/t)$가 볼록함수임을 확인해봅시다. 임의의 $\boldsymbol{x}$, $\boldsymbol{y}$와 0이 아닌 실수 t, s, 그리고 $0\le\theta\le 1$을 만족하는 실수 θ에 대해,

$$
\begin{aligned}
f(\theta \boldsymbol{x}+(1-\theta)\boldsymbol{y}, \theta t+(1-\theta)s) &= (\theta t+(1-\theta)s)h\left(\frac{\theta \boldsymbol{x}+(1-\theta)\boldsymbol{y}}{\theta t+(1-\theta)s}\right) \\
&= (\theta t+(1-\theta)s)h\left(\frac{\theta t(\boldsymbol{x}/t)+(1-\theta)s(\boldsymbol{y}/s)}{\theta t+(1-\theta)s}\right) \\
&= (\theta t+(1-\theta)s)h\left(\frac{\theta t(\boldsymbol{x}/t)+(1-\theta)s(\boldsymbol{y}/s)}{\theta t+(1-\theta)s}\right) \\
&= (\theta t+(1-\theta)s)h\left(\frac{\theta t}{\theta t+(1-\theta)s}(\boldsymbol{x}/t)+\frac{(1-\theta)s}{\theta t+(1-\theta)s}(\boldsymbol{y}/s)\right) \\
&\le (\theta t+(1-\theta)s)\left\{\frac{\theta t}{\theta t+(1-\theta)s}h(\boldsymbol{x}/t)+\frac{(1-\theta)s}{\theta t+(1-\theta)s}h(\boldsymbol{y}/s)\right\} \\
&= \theta t h(\boldsymbol{x}/t)+(1-\theta)s h(\boldsymbol{y}/s) \\
&= \theta f(\boldsymbol{x}, t)+(1-\theta)f(\boldsymbol{y}, s)
\end{aligned}
$$

따라서 함수 f는 볼록함수입니다.

09.

(문제에서 "음이 아닌 확률변수 X"라는 조건이 빠졌습니다!)

$a < 0$ 또는 $a > 1$일 때, $E(X^a) \ge (E(X))^a$

$0 < a < 1$일 때, $E(X^a) \ge (E(X))^a$

10.

(a) $I_c^*(\boldsymbol{y}) = sup_{x \in \mathrm{dom} f}\{\boldsymbol{y}^T\boldsymbol{x} - I_C(\boldsymbol{x})\} = sup_{x \in \mathrm{dom} f}\{\boldsymbol{y}^T\boldsymbol{x}\}$

(b) $f^*(y) = \begin{cases} \sum_{i=1}^{n} y_i \log y_i, & y_i > 0,\ \sum_{i=1}^{n} y_i = 1 \\ \infty, & otherwise \end{cases}$

11.

주어진 이차형식을 합의 형태로 나타내어보면 다음과 같습니다.

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j + \sum_{i=1}^{n} b_i x_i + c$$

① $\underline{\nabla f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}}$

$\boldsymbol{x}^T A\boldsymbol{x}$와 $\boldsymbol{b}^T\boldsymbol{x}$의 편도함수를 구해보면,

$$\begin{aligned} \frac{\partial}{\partial x_k}(\boldsymbol{x}^T A\boldsymbol{x}) = \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j\right) &= \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n}\left(a_{ii}x_i^2 + \sum_{j\neq i} a_{ij}x_i x_j\right)\right) \\ &= 2a_{kk}x_k + \sum_{j\neq k} a_{jk}x_j + \sum_{j\neq k} a_{kj}x_j \\ &= \sum_{j=1}^{n} a_{jk}x_j + \sum_{j=1}^{n} a_{kj}x_j \\ (A\text{가 대칭이므로}) &= 2\sum_{j=1}^{n} a_{kj}x_j \\ \frac{\partial}{\partial x_k}(\boldsymbol{b}^T\boldsymbol{x}) = \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n} b_i x_i\right) &= b_k \end{aligned}$$

따라서 $\frac{\partial}{\partial x_k}(f(\boldsymbol{x})) = \sum_{j=1}^{n} a_{kj}x_j + b_k$이므로 $\nabla f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$입니다.

② $\underline{\nabla^2 f(\boldsymbol{x}) = A}$

①에 의해,

$$\frac{\partial^2}{\partial t\partial k}(f(\boldsymbol{x})) = \frac{\partial}{\partial t}\left(\sum_{j=1}^{n} a_{kj}x_j + b_k\right) = a_{kt}$$

임을 알 수 있습니다. 따라서, $\nabla^2 f(\boldsymbol{x}) = A$입니다.

12.

$g^*(\boldsymbol{y}) = sup_{\boldsymbol{x}}\{(\boldsymbol{y} - \boldsymbol{a})^T\boldsymbol{x} - f(\boldsymbol{x})\} - b = f^*(\boldsymbol{y} - \boldsymbol{a}) - b$

13.

따라서 주어진 볼록 최적화 문제의 해는 다음과 같습니다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{4}\nu_1 - \dfrac{3}{4}\nu_2 \\ -\dfrac{1}{8}\nu_1 + \dfrac{1}{4}\nu_2 \end{bmatrix} = \begin{bmatrix} \dfrac{7}{5} \\ -\dfrac{2}{5} \end{bmatrix}$$

14.

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ 가 볼록함수이므로, 임의의 $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^n$과 $0 \le \theta \le 1$인 실수 θ에 대해

$$f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) \le \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})$$

가 성립합니다. f가 음이 아닌 볼록함수이므로 위 부등식의 양변을 제곱해도 부등식은 변하지 않습니다. 따라서, 위 부등식을 제곱하면

$$\begin{aligned} \{f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})\}^2 &\le \{\theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})\}^2 \\ &= \theta^2 \{f(\boldsymbol{x})\}^2 + 2\theta(1-\theta) f(\boldsymbol{x}) f(\boldsymbol{y}) + (1-\theta)^2 \{f(\boldsymbol{y})\}^2 \end{aligned}$$

을 얻습니다. 이 부등식의 양변에 $\theta\{f(\boldsymbol{x})\}^2$과 $(1-\theta)\{f(\boldsymbol{y})\}^2$을 각각 더하고 빼면,

$$\begin{aligned} \{f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})\}^2 &\le \theta^2 \{f(\boldsymbol{x})\}^2 + 2\theta(1-\theta) f(\boldsymbol{x}) f(\boldsymbol{y}) + (1-\theta)^2 \{f(\boldsymbol{y})\}^2 \\ &\quad + \theta\{f(\boldsymbol{x})\}^2 - \theta\{f(\boldsymbol{x})\}^2 + (1-\theta)\{f(\boldsymbol{y})\}^2 - (1-\theta)\{f(\boldsymbol{y})\}^2 \\ &= \theta\{f(\boldsymbol{x})\}^2 + (1-\theta)\{f(\boldsymbol{y})\}^2 - \theta(1-\theta)\{f(\boldsymbol{x}) - f(\boldsymbol{y})\}^2 \\ &\le \theta\{f(\boldsymbol{x})\}^2 + (1-\theta)\{f(\boldsymbol{y})\}^2 \end{aligned}$$

이므로, f^2도 볼록함수입니다.

15.

간단한 코딩이므로 직접 해봅시다!

16.

(a) [정리 3-2]의 이계조건을 이용하기 위해 최소자승법의 목적함수

$$f(a_0, a_1) = \frac{1}{N} \sum_{i=1}^{N} (a_0 + a_1 x_i - y_i)^2$$

의 헤세 행렬을 구해봅시다.

$$\nabla^2 f(a_0,a_1)=\begin{pmatrix}\dfrac{\partial^2 f}{\partial a_0^2} & \dfrac{\partial^2 f}{\partial a_0 \partial a_1}\\ \dfrac{\partial^2 f}{\partial a_1 \partial a_0} & \dfrac{\partial^2 f}{\partial a_1^2}\end{pmatrix}=\begin{pmatrix}2 & \dfrac{2}{N}\sum_{i=1}^{N}x_i\\ \dfrac{2}{N}\sum_{i=1}^{N}x_i & \dfrac{2}{N}\sum_{i=1}^{N}x_i^2\end{pmatrix}=2\begin{pmatrix}1 & E(X)\\ E(X) & E(X^2)\end{pmatrix}$$[1]

$\mathbb{R}^2$의 임의의 원소 $(a\ b)$에 대해

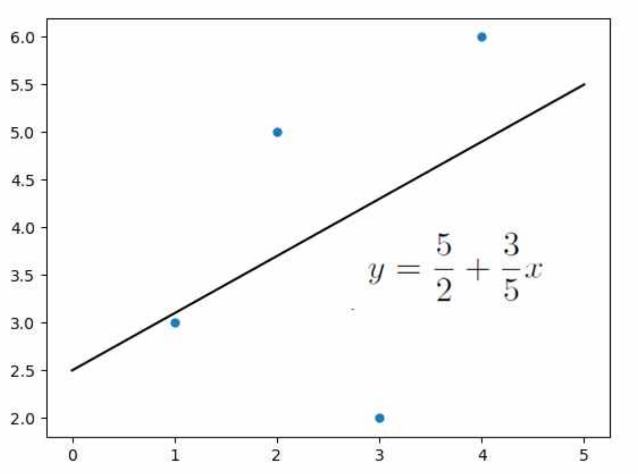
$$\begin{aligned}(a\ b)\nabla^2 f(a_0,a_1)\begin{pmatrix}a\\b\end{pmatrix} &= 2(a\ b)\begin{pmatrix}1 & E(X)\\ E(X) & E(X^2)\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}\\ &= 2\left(a^2+2E(X)ab+E(X^2)b^2\right)\\ &= 2\left((a+E(X)b)^2+\left(E(X^2)-(E(X))^2\right)b^2\right)\end{aligned}$$

이고, $E(X^2)-(E(X))^2=V(X)\geq 0$을 이용하면

$$(a\ b)\nabla^2 f(a_0,a_1)\begin{pmatrix}a\\b\end{pmatrix}=2\left((a+E(X)b)^2+\left(E(X^2)-(E(X))^2\right)b^2\right)\geq 0$$

임을 알 수 있습니다. 따라서, 함수 f의 해세 행렬 $\nabla^2 f(a_0,a_1)$이 준정부호 행렬이므로 최소자승법의 목적함수 f는 볼록함수입니다.

(b) 주어진 점들의 추세를 가장 잘 표현하는 직선의 방정식은 $y=\dfrac{5}{2}+\dfrac{3}{5}x$입니다.



---

[1] 여기서 $E(X)=\dfrac{1}{N}\sum_{i=1}^{N}x_i$, $E(X^2)=\dfrac{1}{N}\sum_{i=1}^{N}x_i^2$입니다.