Exercise 07

Friday, November 24, 2023 2:22 PM

Exercise 25 (8 points)

Consider the multi-objective optimization problem

$$\begin{aligned} &\text{minimize}_x && f_1(x), \ldots, f_\ell(x) \\ &\text{subject to} && g_i(x) \le 0 \ (i = 1 \ldots m) \\ & && h_j(x) = 0 \ (j = 1 \ldots p) \end{aligned}$$

Suppose that the f_k, g_i are differentiable and convex and that the h_j are affine. For a point $c \in \mathbb{R}^\ell$ we can form the single objective function $f_c(x) := c_1 f_1(x) + \ldots + c_\ell f_\ell(x)$

Suppose that $x^* \in \mathbb{R}^n$, $c^* \in \mathbb{R}^\ell$, $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ satisfy the following conditions:

$$\begin{aligned} \lambda_i^* &\ge 0 \\ c_k^* &\ge 0 \\ \nabla f_{c^*}(x^*) + \textstyle\sum_i \lambda_i^* \nabla g_i(x^*) + \sum_j \mu_j^* \nabla h_j(x^*) &= 0 \\ \textstyle\sum_i \lambda_i^* g_i(x^*) &= 0 \\ g_i(x^*) &\le 0 \\ h_j(x^*) &= 0 \end{aligned}$$

Then show that x^* is a Pareto optimal point for the multiobjective problem.

Since $f_k(x)$, $k=1,\ldots,\ell$ are convex and $c_k^* \ge 0$

The positive linear combination $f_{c^*}(x) = \sum_{k=1}^{\ell} c_k^* f_k(x)$ is convex

Therefore, the following problem is a convex optimization problem

$$\begin{aligned} &\text{minimize} && f_{c^*}(x) \\ &\text{subject to} && g_i(x) \le 0 \quad (i = 1, \ldots, m) \\ & && h_j(x) = 0 \quad (j = 1, \ldots, p) \end{aligned}$$

Furthermore, since $(x^*, (\lambda^*, \mu^*))$ satisfies KKT condition, x^* is an optimal point, $p^* = f_{c^*}(x^*)$

$\Rightarrow \forall x \in \operatorname{dom} f_k \cap \operatorname{dom} g_i \cap \operatorname{dom} h_j$,

$$f_{c^*}(x^*) \le f_{c^*}(x)$$

Suppose x^* is not a Pareto optimal point of the multiobjective problem.

then $\exists \tilde{x} \in \operatorname{dom} f_k \cap \operatorname{dom} g_i \cap \operatorname{dom} h_j$, s.t.

for $\forall k = 1, \ldots, \ell$, $f_k(\tilde{x}) \le f_k(x^*)$ ①, and

strictly

$\exists \tilde{k} \in \{1, \ldots, \ell\}$, $f_{\tilde{k}}(\tilde{x}) \ne f_{\tilde{k}}(x^*)$, meaning $f_{\tilde{k}}(\tilde{x}) < f_{\tilde{k}}(x^*)$ ②

$$\begin{aligned} \text{then } f_{c^*}(\tilde{x}) &= \sum_{k=1}^{\ell} c_k^* f_k(\tilde{x}) \\ &= c_{\tilde{k}}^* f_{\tilde{k}}(\tilde{x}) + \sum_{\substack{k=1,\\ k\ne\tilde{k}}}^{\ell} c_k^* f_k(\tilde{x}) \\ &\overset{\text{by ①}}{\le} c_{\tilde{k}}^* f_{\tilde{k}}(\tilde{x}) + \sum_{\substack{k=1\\ k\ne\tilde{k}}}^{\ell} c_k^* f_k(x^*) \\ &\overset{\text{by ②}}{<} c_{\tilde{k}}^* f_{\tilde{k}}(x^*) + \sum_{\substack{k=1\\ k\ne\tilde{k}}}^{\ell} c_k^* f_k(x^*) = \sum_{k=1}^{\ell} c_k^* f_k(x^*) = f_{c^*}(x^*) \end{aligned}$$

$f_{c^*}(\tilde{x}) < f_{c^*}(x^*)$ $\perp$ contradiction!

$\Rightarrow x^*$ is a Pareto optimal point for the multiobjective problem.

Exercise 27 (10 points)

(a) Consider an $\mathbb{R}^3$-valued random variable (X_1, X_2, X_3) with density function $f(x_1, x_2, x_3) := (x_1^2 + x_2^2 + x_3^2)\chi_I$, where χ_I is the indicator function of the unit cube $I := [0,1]^3$, i.e. $\chi_I(x) = 1$ if $x \in I$ and $= 0$ otherwise.

(i) (2 points) Compute the probability $P(X_1 \le \frac{1}{2}, X_3 \ge \frac{1}{2})$

(ii) (2 points) Compute the density function of the $\mathbb{R}^2$-valued random variable (X_1, X_2).

(b) (2 points) Consider the $\mathbb{R}^2$-valued random variable (X, Y) with density function $f(x, y) := e^{-x-y}$ for $x, y \ge 0$ and 0 otherwise. Are the random variables X and Y independent?

(c)

(i) (2 points) Consider two independent random variables X and Y taking values 1 or -1 each with probability $\frac{1}{2}$. Let $Z := X \cdot Y$. Show that X and Z are independent, and that Y and Z are independent.

(ii) (2 points) Show that X, Y and Z are not jointly independent, in the sense that $P(X = a, Y = b, Z = c)$ is not always equal to $P(X = a) \cdot P(Y = b) \cdot P(Z = c)$.

(a) (i) $P(X_1 \le \frac{1}{2}, X_3 \ge \frac{1}{2}) = \int_{-\infty}^{\frac{1}{2}} \int_{-\infty}^{+\infty} \int_{\frac{1}{2}}^{+\infty} (x_1^2 + x_2^2 + x_3^2) \chi_I \, dx_1 \, dx_2 \, dx_3$

$$= \int_0^{\frac{1}{2}} \int_0^1 \int_{\frac{1}{2}}^1 (x_1^2 + x_2^2 + x_3^2) \, dx_1 \, dx_2 \, dx_3$$

$$= \frac{1}{2} \int_0^{\frac{1}{2}} x_1^2 dx_1 + \frac{1}{4} \int_0^1 x_2^2 dx_2 + \frac{1}{2} \int_{\frac{1}{2}}^1 x_3^2 dx_3$$

$$= \frac{1}{2} \cdot \frac{1}{3} \left(\frac{1}{2}\right)^3 + \frac{1}{4} \cdot \frac{1}{3} \cdot (1)^3 + \frac{1}{2} \cdot \frac{1}{3} \left(1^3 - \left(\frac{1}{2}\right)^3\right)$$

$$= \frac{1}{4}$$

$\frac{1}{16} + \frac{4}{16} + \frac{7}{16}$

$\frac{12}{16} = \frac{3}{4}$

(ii) $f(x_1, x_2) = \int_{-\infty}^{+\infty} (x_1^2 + x_2^2 + x_3^2) \chi_I \, dx_3$

$$= \int_0^1 (x_1^2 + x_2^2 + x_3^2) \, dx_3$$

$$= x_1^2 + x_2^2 + \frac{1}{3} \cdot 1^3 = x_1^2 + x_2^2 + \frac{1}{3}$$

(b) Suppose $a, b, c, d \in \mathbb{R}$ are arbitrary, and $b \ge a$, $d \ge c$

Let $u(t) = \begin{cases} 1, & t \ge 0 \\ 0, & t < 0 \end{cases}$ $\Rightarrow$ $f(x,y) = u(x)u(y)e^{-x-y}$

$P(a \le x \le b, c \le y \le d) = \int_a^b \int_c^d u(x)u(y) e^{-x-y} dx dy$

$$= \int_a^b \int_c^d u(x) e^{-x} \cdot \underbrace{u(y) e^{-y}}_{\text{independent of } x} dx dy$$

$$= \int_a^b u(y) e^{-y} \underbrace{\int_c^d u(x) e^{-x} dx}_{\text{ind. of } y} dy$$

$$= \int_c^d u(x) e^{-x} dx \cdot \int_c^d u(y) e^{-y} dy$$

$P(a \le X \le b) = \int_a^b \int_{-\infty}^{+\infty} u(x)u(y) e^{-x-y} dy \, dx$

$$= \int_a^b u(x) e^{-x} dx \cdot \left(\int_0^{+\infty} e^{-y} dy + \int_{-\infty}^0 0 \cdot e^{-y} dy\right)$$

$$= \int_a^b u(x) e^{-x} dx \cdot \underbrace{-e^{-y}\Big|_0^{+\infty}}_{= (0 - (-1)) = 1}$$

$$= \int_a^b u(x) e^{-x} dx$$

Similarly,

$P(c \le Y \le d) = \int_c^d u(y) e^{-y} dx$

Therefore $\forall a, b, c, d \in \mathbb{R}$, $b \ge a$, $d \ge c$

$$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b) \cdot P(c \le Y \le d)$$

$\Rightarrow$ X and Y are independent

(c) (i) X, Y, Z are discrete random variables

$X : \Omega \to \{-1, 1\}$

$Y : \Omega \to \{-1, 1\}$

$\Rightarrow$ $Z = X \cdot Y : \Omega \to \{-1, 1\}$

$P(Z = -1) = P(X \cdot Y = -1) = P((X=1, Y=-1) \cup (X=-1, Y=1))$

$$= P(X=1, Y=-1) + P(X=-1, Y=1)$$

$$\underset{\text{X, Y ind.}}{=} P(X=1) \cdot P(Y=-1) + P(X=-1) P(Y=1)$$

$$= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

Similarly,

$P(Z = 1) = P(X \cdot Y = 1) = P(X=1) P(Y=1) + P(X=-1) P(Y=-1)$

$$= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

① Prove that X and Z are independent

Case 1: $X = 1$, $Z = 1$

$P(X=1, Z=1) = P(X=1, XY=1)$

$= P(X=1, Y=1)$

$\overset{X,Y\ ind.}{\Longrightarrow} P(X=1)P(Y=1) = \frac{1}{4} = P(X=1)P(Z=1)$ ✓

Case 2: $X=1,\ Z=-1$

$P(X=1, Z=-1) = P(X=1, Y=-1)$

$= P(X=1)P(Y=-1) = \frac{1}{4} = P(X=1)P(Z=-1)$ ✓

Similarly: Case 3: $X=-1,\ Z=1$

$P(X=-1, Z=1) = P(X=-1)P(Y=-1)$

$= \frac{1}{4} = P(X=-1)P(Z=1)$ ✓

Case 4: $X=-1,\ Z=-1$

$P(X=-1, Z=-1) = P(X=-1)P(Y=1)$

$= \frac{1}{4} = P(X=-1)P(Z=-1)$ ✓

All the elementary events are independent

$\Rightarrow$ X and Z are independent. $\square$

② $\underline{X \cdot Z} = X \cdot X \cdot Y = X^2 Y = \underline{Y}$

By ① we proved for independent X and Z with described prob. distribution, X and $X \cdot Z = Y$ are independent $\square$

(ii) Let $a=b=c=1$

$P(X=1, Y=1, Z=1)$

$= P(X=1, Y=1, XY=1)$

$= P(X=1, Y=1)$

$= P(X=1) \cdot P(Y=1) = \frac{1}{4}$

$P(X=1) \cdot P(Y=1) \cdot P(Z=1)$

$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \neq P(X=1, Y=1, Z=1)$

$\Rightarrow$ X, Y and Z are not jointly independent.

Exercise 28 (12 points)

In the following all random variables are $\mathbb{R}^n$-valued and should be defined on the same fixed probability space (Ω, P). That is: a random variable is a map $\Omega \to \mathbb{R}^n$.

Two random variables X, Y are called *independent*, if they satisfy $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$ for all $A, B \subseteq \mathbb{R}^n$

One can add and scalar multiply random variables according to the rules $(X+Y)(\omega) := X(\omega) + Y(\omega)$ and $(\lambda X)(\omega) := \lambda \cdot X(\omega)$. With this, random variables on (Ω, P) form a vector space, so the notion of *linear independence* makes sense. In this exercise you should explore the relationship between the linear independence and the independence of random variables.

An $\mathbb{R}^n$-valued random variable Z is called *almost surely constant*, if there is a $v \in \mathbb{R}^n$ such that $P(Z = v) = 1$.

Let X, Y, Z be $\mathbb{R}$-valued random variables.

(a) Show that $\mathrm{Var}(Z) = 0$ if and only if Z is almost surely constant.

"$\Rightarrow$" Suppose $Var(Z) = 0$, $Z = (Z_1, \cdots, Z_n)^T$

$E((Z - EZ)(Z - EZ)^T) = 0$

for $\forall i \in \{1, \dots, n\}$

$E((Z_i - EZ_i)^2) = 0 \Rightarrow \sum_{\omega \in \Omega} (Z_i - EZ_i)^2(\omega) P(\omega) = 0$

Since $(Z_i - EZ_i)^2(\omega) \geq 0$, $P(\omega) \geq 0$, $\forall \omega \in \Omega$

$\Rightarrow (Z_i - EZ_i)(\omega) = 0$, for $\forall \omega \in \Omega$

$\Rightarrow P(Z_i - EZ_i = 0) = 1$

$\Rightarrow P(Z_i = EZ_i) = 1$ for $\forall i \in \{1, \dots, n\}$

$\Rightarrow P(Z = EZ) = 1$

$\exists v = EZ$, s.t. $P(Z = v) = 1$

"$\Leftarrow$" Suppose $\exists v \in \mathbb{R}^n$ s.t. $P(Z = v) = 1$,

$EZ = v \cdot P(Z = v) = v$

$Var(Z) = E((Z - EZ)(Z - EZ)^T)$

$= E(ZZ^T) - E(Z \cdot (EZ)^T) - E((EZ) \cdot Z^T) + EZ\,(EZ)^T$

$= E(ZZ^T) - EZ \cdot (EZ)^T$

$= P(ZZ^T = vv^T) \cdot vv^T - v \cdot v^T$

$E\begin{pmatrix} EZ_1 \cdot Z_1 & & \\ & \ddots & \\ & & EZ_n \cdot Z_n \end{pmatrix} = \begin{pmatrix} EZ \cdot \Sigma r P & & \\ & \ddots & \\ & & EZ \cdot r P \end{pmatrix}$

$= \begin{pmatrix} EZ \cdot EZ & & \\ & \ddots & \\ & & \end{pmatrix}$

Too complicated

$\Rightarrow \{Var(Z)\}_{ij} = E((Z_i - EZ_i)(Z_j - EZ_j))$, $i, j = 1, \dots, n$

for $i = j$,

$\{Var(Z)\}_{ii} = \sum_{\omega \in \Omega} (Z_i - EZ_i)^2(\omega) \cdot P(\omega) = 0$, with $\sum_{\omega \in \Omega} P(\omega) = 1$

- suppose for $\forall w_i = (Z_i - EZ_i)^2(\omega) \geq 0$, $\omega \in \Omega$
  s.t. $P((Z_i - EZ_i)^2 = w_i) \neq 1$, therefore < 1

then $\exists j, k \in \{1, \dots, n\}$,

$\Rightarrow tr(E(WW^T)) = \sum_{i=1}^{n} \sum_{\omega \in \Omega} W_i^2(\omega) P(\omega) = 0$

$= \sum_{\omega \in \Omega} \sum_{i=1}^{n} W_i^2(\omega) P(\omega)$

$= \sum_{\omega \in \Omega} (W^T W)(\omega) P(\omega) = 0 \quad (*)$

- suppose for $\forall\ w \in W^T W(\omega) \geq 0$, $\omega \in \Omega$
  s.t. $P(W^T W = w) \neq 1$, therefore $\boxed{<1}$

Since $\sum_{\omega \in \Omega} P(\omega) = 1$

then $\exists j \in \{1, \dots, n\}$, s.t. $0 < P(W^T W = w_j) < 1$

and $\exists k \in \{1, \dots, n\}$, $k \neq j$ s.t. $0 < P(W^T W = w_k) \leq 1 - P(W^T W = w_j)$

$(*)$ $\sum_{\omega \in \Omega} (W^T W)(\omega) P(\omega) = w_j P(W^T W = w_j) + w_k P(W^T W = w_k) + \overset{\text{(the rest)}}{\sum_{\substack{\omega \in \Omega \\ W^T W \neq w_k, w_j}} (W^T W)(\omega) P(\omega)}$

$w_j, w_k \geq 0$, $w_j \neq w_k$ $\Rightarrow$

$w_j P(W^T W = w_j) + w_k P(W^T W = w_k) > 0 \Rightarrow Var(Z) > 0$ $\perp$ contradition!

$\Rightarrow \exists w \in W^T W(\omega) \geq 0$, s.t. $P(W^T W = w) = 1$

$\Rightarrow \exists w \in \|Z - EZ\|^2 \Longrightarrow \dots$

$$= E(ZZ^T) - E(Z\cdot(EZ)^T) - E((EZ)\cdot Z^T) + EZ\,(EZ)^T$$

$$= \left(\quad \cdots \quad \right)$$

$$= E(ZZ^T) - EZ\cdot(EZ)^T$$

$$= \underbrace{P(ZZ^T = vv^T)}_{P(Z=v)}\cdot vv^T - v\cdot v^T$$

$$= vv^T - vv^T = 0 \quad \square$$

$\Rightarrow \exists\, w \in W^T W(\omega) \geq 0,$ s.t. $P(W^T W = w) = 1$

$\Rightarrow \exists\, w \in \|Z - EZ\|^2 \Longrightarrow \cdots$

(b) Show that the almost surely constant random variables form a subvector space of all random variables.

Suppose the subset U is the set of all almost surely constant r.v.s

"0" $\forall \omega \in \Omega$, $0(\omega) = 0 \quad \Rightarrow P(0(\omega) = 0) = 1$

$\Rightarrow 0(\omega) \in U$

Additivity Suppose $X, Y \in U$ are arbitrary

Then $\exists\, v_1, v_2 \in \mathbb{R}^n$, s.t. $P(X = v_1) = 1$, $P(Y = v_2) = 1$

$$1 \geq P(X + Y = v_1 + v_2) \geq P(X = v_1, Y = v_2)$$

$$\geq P(X = v_1)P(Y = v_2) = 1$$

$1 = 1 \Rightarrow$ the inequalities become equalities

$\Rightarrow \exists\, v = v_1 + v_2 \in \mathbb{R}^n$, s.t. $P(X + Y = v) = 1$

$\Rightarrow (X + Y)(\omega) \in U$

Scaling Suppose $X \in U$, $\lambda \in \mathbb{R}$ are arbitrary

Then $\exists\, v_1 \in \mathbb{R}^n$, s.t. $P(X = v_1) = 1$

If $\lambda \neq 0$, $P(\lambda X = \lambda v_1) = P(X = v_1) = 1$

$\exists\, v = \lambda v_1 \in \mathbb{R}^n$, s.t. $P(\lambda X = v) = 1 \Rightarrow \lambda X(\omega) \in U$

If $\lambda = 0$, $\lambda X(\omega) = 0(\omega) \in U$

Therefore U is a subvector space of all random variables.

(c) Show that if Z is not almost surely constant and Z and Y are linearly dependent, then their correlation coefficient is 1 or -1

Z and Y are linearly dependent

$\Rightarrow Z = \lambda Y \quad \lambda \neq 0, \lambda \in \mathbb{R}$ since Z is not almost surely constant

Here we suppose Z and Y are $\mathbb{R}$-valued random variables

$$Cov(Z, Y) = Cov(\lambda Y, Y)$$

$$\overset{\text{bilinearity}}{=} \lambda\, Cov(Y, Y) = \lambda\, \overbrace{Var(Y)}^{>0}$$

$$\sigma_Z = \sqrt{Cov(Z, Z)} = \sqrt{Cov(\lambda Y, \lambda Y)}$$

$$= \sqrt{\lambda^2 Cov(Y, Y)}$$

$$= |\lambda|\, \sigma_Y$$

So the correlation coefficient $\rho_{ZY} = \dfrac{Cov(Z, Y)}{\sigma_Z\, \sigma_Y}$

$$= \frac{\lambda\, \sigma_Y^2}{|\lambda|\, \sigma_Y^2} = \frac{\lambda}{|\lambda|}$$

$$\Rightarrow \rho_{ZY} = \begin{cases} 1, & \text{if } \lambda > 0 \\ -1, & \text{if } \lambda < 0 \end{cases}$$

(d) Does linear independence of two random variables X, Y imply that they are independent?

No. For instance $Y = X^2$, with $P(X = 2) = \frac{1}{2}$, $P(X = 3) = \frac{1}{2}$

$\Rightarrow \forall \lambda \in \mathbb{R}$, $\lambda X \neq Y$

$P(Y = 4) = \frac{1}{2}$, $P(Y = 9) = \frac{1}{2}$

$P(Y = 4, X = 2) = P(X = 2) = \frac{1}{2} \neq P(Y = 4)\cdot P(X = 2)$

(e) Does independence of two random variables X, Y imply that they are linearly independent?

(f) Does the answer in part (e) change, if one assumes X, Y to be not almost surely constant?

[2 points per item]

If X, Y are not almost surely constant, Yes

Suppose X and Y are independent

$X: \Omega \to T$, $Y: \Omega \to S$

$\Rightarrow$ For $\forall A \in T$, $B \in S$

$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$

Suppose X and Y are linearly dependent

$\exists \lambda \in \mathbb{R}, \lambda \neq 0$, $X = \lambda Y$,

$\exists A \in T$, $0 < P(x \in A) < 1$

let $B = \{Y(\omega) \mid \omega \in \Omega, X(\omega) \in A\} = \{\lambda X(\omega) \mid \omega \in \Omega\}$

$X \in A \iff Y \in B$

so $P(x \in A) = P(y \in B)$

$P(x \in A, Y \in B) = P(x \in A) \neq P(x \in A) \cdot P(y \in B)$ $\perp$ contradiction!

$P(x \in A) \neq 0$

$P(x \in A) \neq 1$

If X and Y are almost surely constant, not necessary

For $v_1, 2v_1$, where $P(X = v_1) = 1$ and $P(Y = 2v_1) = 1$

$X = 2Y$, X and Y are linearly dependent

$P(X = v_1, Y = 2v_1) = P(X = v_1) P(Y = 2v_1)$, independent!