به نام خدا

پاسخ تمرین سری **سوم** درس بهینه‌سازی

(نیمسال اول ۱۴۰۲)

۱- برنامه‌ریزی خطی زیر را در نظر بگیرید:

$$\textit{minimize } 4x_1 + 3x_2$$
$$\textit{subject to } 5x_1 + x_2 \geq 11$$
$$2x_1 + x_2 \geq 8$$
$$x_1 + 2x_2 \geq 7$$
$$x_1, x_2 \geq 0$$

دوگان متناظر با این مسأله را بنویسید و پاسخ مسأله دوگان را بیابید.

پاسخ:

$$\textit{maximize } 11\lambda_1 + 8\lambda_2 + 7\lambda_3$$
$$\textit{subject to } 5\lambda_1 + 2\lambda_2 + \lambda_3 \leq 4$$
$$\lambda_1 + \lambda_2 + 2\lambda_3 \leq 3$$
$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$

پاسخ مسأله دوگان عبارت است از:

$$\lambda^T = c_B^T B^{-1} = \left[0, \frac{5}{3}, \frac{2}{3}\right]$$

۲- فرض کنید یک سیستم فرستنده بیسیم دارای n فرستنده است. فرستنده j با توان $p_j \geq 0$ ارسال را انجام می‌دهد و m موقعیت مختلف برای دریافت سیگنال‌ها در نظر گرفته شده است. بهره فرستنده j به موقعیت i را با $g_{i,j}$ نشان می‌دهیم. بنابراین توان سیگنال ارسال شده توسط فرستنده j و دریافت شده در موقعیت i برابر است با $g_{i,j}p_j$ و در نتیجه توان کل دریافتی در موقعیت i برابر حاصل جمع تمامی توان‌های دریافتی از همه فرستنده‌ها است. می‌خواهیم مینیمم مجموع توان ارسالی را بیابیم به این شرط که توان دریافتی در هر موقعیت حداقل P باشد.

الف) این مسأله را فرمول‌بندی کنید.

پاسخ: تابع هدف عبارت است از $p_1+\cdots+p_n$. قید مربوط به موقعیت i-ام عبارت است از $g_{i,1}p_1+\cdots+g_{i,n}p_n \geq P$. بنابراین مسأله بهینه‌سازی مربوطه عبارت است از:

$$\begin{aligned} & minimize\ p_1+\cdots+p_n \\ & subject\ to\ g_{i,1}p_1+\cdots+g_{i,n}p_n \geq P, \qquad i=1,\dots,m \\ & p_1,\dots,p_n \geq 0 \end{aligned}$$

با تعریف ماتریس G و بردار p می‌توانیم چنین بازنویسی کنیم:

$$\begin{aligned} & minimize\ 1^T p \\ & subject\ to\ Gp \succcurlyeq P \\ & p \succcurlyeq 0 \end{aligned}$$

ب) شرایط KKT را برای آن بنویسید.

پاسخ:

$$\begin{aligned} & \mathbf{1}^T-\lambda^T G-\nu^T=\mathbf{0}^T \\ & \lambda^T(P\mathbf{1}-Gp)-\nu^T p=0 \\ & Gp \geq P\mathbf{1} \\ & \lambda,\nu,p \geq 0 \end{aligned}$$

۳- فرض کنید x_i که در آن $i=1,\dots,n$ متغیرهای تصادفی مستقل از هم باشند که از توزیع پواسون پیروی می‌کنند طوری‌که:

$$prob(x_i=k)=\frac{e^{-\mu_i}\mu_i^k}{k!},$$

باشد که در آن میانگین‌های μ_i نامعلوم هستند. متغیرهای x_i تعداد دفعاتی را نشان می‌دهند که یکی از n رویداد مستقل در یک دوره معین رخ داده است. هدفی که دنبال می‌شود آن است که بتوانیم میانگین‌های μ_i را با استفاده از تعداد m آشکارساز تخمین بزنیم. در آزمایشی که بدین منظور انجام می‌شود اگر رویداد i رخ دهد، احتمال آن‌که آشکارساز j بتواند آن را تشخیص دهد p_{ji} خواهد بود. فرض می‌کنیم که مقادیر احتمال‌های p_{ji} داده شده باشند طوری‌که برای آن‌ها داریم: $p_{ji} \geq 0$ و $\sum_{j=1}^m p_{ji} \leq 1$. تعداد کل رویدادهای ثبت شده توسط آشکارساز j با y_j نشان داده می‌شود بنابراین:

$$y_j=\sum_{i=1}^n y_{ji}\,, j=1,\dots,m.$$

مساله‌ی تخمین بیشینه‌درستنمایی[1] را برای تخمین میانگین‌های μ_i براساس مشاهدات $y_j, j=1,\dots,m$، به‌صورت یک مساله بهینه‌سازی محدب بنویسید.

(راهنمایی: متغیرهای y_{ji} دارای توزیع پواسون با میانگین $p_{ji}\mu_i$ هستند به این معنی که برای آن‌ها داریم:

$$prob(y_{ji}=k)=\frac{e^{-p_{ji}\mu_i}(p_{ji}\mu_i)^k}{k!}.$$

مجموع n متغیر تصادفی مستقل با توزیع پواسون با میانگین‌های $\lambda_1,\dots,\lambda_n$ دارای توزیع پواسون با میانگین $\lambda_1+\cdots+\lambda_n$ است.

پاسخ: با توجه به آنچه که در صورت مساله بیان شده است مقادیر y_jها برای $j=1,\dots,m$ آشکارساز مشاهده شده است. براساس راهنمایی مساله می‌دانیم y_j دارای توزیع پواسون با میانگین

$$\sum_{i=1}^{n} p_{ji}\mu_i = p_j^T\mu$$

است. حال اگر فرض کنیم آشکارساز y_j تعداد k_j رویداد را تشخیص داده باشد داریم:

$$LF=P(y_1=k_1)\times P(y_2=k_2)\times\dots\times P(y_m=k_m)=\prod_{j=1}^{m}P(y_j=k_j)$$
$$=\prod_{j=1}^{m}\frac{\left(e^{-\sum_{i=1}^{n}p_{ji}\mu_i}\left(\sum_{i=1}^{n}p_{ji}\mu_i\right)^{k_j}\right)}{k_j!}$$

بنابراین با گرفتن لگاریتم از دوطرف داریم:

$$LLF=\sum_{j=1}^{m}\left(-\sum_{i=1}^{n}p_{ji}\mu_i+k_j\log\left(\sum_{i=1}^{n}p_{ji}\mu_i\right)-\log k_j!\right)$$
$$=-\sum_{j=1}^{m}p_j^T u+\sum_{j=1}^{m}k_j\log(p_j^T u)-\sum_{j=1}^{m}\log k_j!$$

توجه داشته باشید که عبارت $\sum_{j=1}^{m}\log k_j!$ دارای مقدار ثابت است بنابراین مساله‌ی بهینه‌سازی تخمین ML به‌صورت زیر نوشته می‌شود:

[1] Maximum likelihood

$$\begin{gathered} maximize \quad -\sum_{j=1}^{m} p_j^T u + \sum_{j=1}^{m} k_j \log\left(p_j^T u\right) \\ \mu \succcurlyeq 0 \end{gathered}$$

۴- فرض کنید $y \in \{0,1\}$ یک متغیر تصادفی باشد که به‌صورت زیر تعریف شده است:

$$y = \begin{cases} 1 & a^T u + b + v \leq 0 \\ 0 & a^T u + b + v > 0 \end{cases}$$

در این رابطه $u \in \mathbb{R}^n$ است و v نیز یک متغیر تصادفی با توزیع گوسی با میانگین صفر و واریانس یک است.

تخمین ML را برای یافتن a و b، با فرض آنکه داده‌های $(u_i, y_i), i = 1, \dots, N$ داده شده باشند به‌صورت یک مساله‌ی بهینه‌سازی محدب بنویسید.

پاسخ: مشاهدات (u_i, y_i) را طوری مرتب می‌کنیم که N_1 تای اول $y = 1$ باشد و بقیه $y = 0$ باشند. بدین‌ترتیب تابع درستنمایی مطابق با رابطه‌ی زیر نوشته می‌شود:

$$lf = \prod_{i=1}^{N_1} P(a^T u_i + b + v \leq 0) \prod_{i=N_1+1}^{N} P(a^T u_i + b + v > 0)$$

حال می‌دانیم $v \sim N(0,1)$ بنابراین:

$$z = a^T u_i + b + v \sim N(a^T u_i + b, 1)$$

حال با فرض

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$$

داریم:

$$P(z \leq 0) = G\left(\frac{0 - (a^T u_i + b)}{1}\right)$$

$$P(z > 0) = 1 - G\big(0 - (a^T u_i + b)\big)$$

$$LLF = \sum_{i=1}^{N_1} \log\Big(G\big(-(a^T u_i + b)\big)\Big) + \sum_{i=N_1+1}^{N} \log\big(1 - G(-a^T u_i - b)\big)$$

همانطورکه می‌دانیم این تابع نسبت به a و b یک تابع مقعر است.

۵- مساله تقریب زیر را درنظر بگیرید.

$$\min \sum_{i=1}^{m} \phi(r_i)$$

$$s.t \quad r = Ax - b$$

$$x \in \mathbb{R}^n, \qquad A \in \mathbb{R}^{m\times n}$$

a) دوگان این مساله را برحسب تابع conjugate جریمه بیابید.

پاسخ: برای محاسبه‌ی دوگان تابع فوق ابتدا به محاسبه‌ی تابع لاگرانژین می‌پردازیم و سپس دوگان را با استفاده از آن محاسبه می‌کنیم؛ داریم:

$$\mathcal{L}(x,r,\nu) = \sum_{i=1}^{m} \phi(r_i) + \nu^T(Ax - b - r)$$

$$\begin{aligned} g(\nu) &= \min_{x,r}\big(\mathcal{L}(x,r,\nu)\big) \\ &= \min_{r}\left(\min_{x}\left(\sum_{i=1}^{m} \phi(r_i) - \nu^T b - \nu^T r + A^T \nu x\right)\right) \\ &= \min_{r}\left(\begin{cases} \sum_{i=1}^{m} \phi(r_i) - \nu^T b - \nu^T r & A^T\nu = 0 \\ -\infty & A^T\nu \neq 0 \end{cases}\right) \end{aligned}$$

حال درصورتی که فرض شود شرط $A^T\nu = 0$ برقرار باشد داریم:

$$g(\nu) = \min\left(\sum_{i=1}^{m} \phi(r_i) - \sum_{i=1}^{m} \nu_i r_i - b^T\nu\right) = \min \sum_{i=1}^{m} (\phi(r_i) - \nu_i r_i) - b^T\nu$$

حال با به‌کارگیری تعریف تابع کانجوگیت داریم:

$$g(\nu) = \sum_{i=1}^{m} -\sup\big(\nu_i r_i - \phi(r_i)\big) - b^T\nu = \sum_{i=1}^{m} -\phi^*(\nu_i) - b^T\nu$$

بنابراین مساله دوگان به‌صورت زیر نوشته می‌شود:

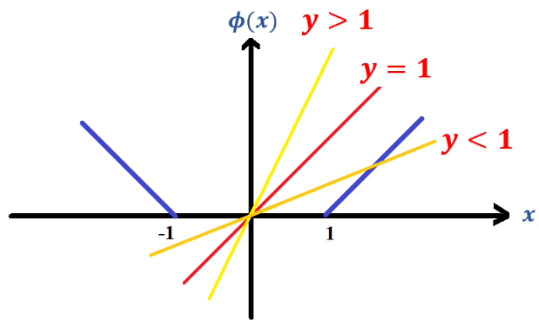
$$\begin{aligned} max \quad & \sum_{i=1}^{m} -\phi^*(\nu_i) - b^T\nu \\ s.t \quad & A^T\nu = 0 \end{aligned}$$

b) اگر تابع جریمه $\phi(x)$، deadzone linear با پهنای $w = 1$ باشد، دوگان مساله‌ی بالا را بیابید.

پاسخ: با توجه به آنکه در این بخش تابع $\phi(x)$ داده شده است و مساله‌ی دوگان به‌دست آمده در بخش قبل براساس برحسب تابع کانجوگیت $\phi(x)$ است، کافی است $\phi^*(x)$ را بیابیم و در مساله‌ی دوگان بخش قبل جایگذاری کنیم؛ داریم:

$$\phi(x) = \begin{cases} 0 & |x| \le 1 \\ |x| - 1 & |x| > 1 \end{cases}$$

$$\phi^*(y) = \sup\big(yx - \phi(x)\big)$$



با توجه به شکل فوق مشخص است $\phi^*(y)$ که در واقع max فاصله‌ی خط yx با $\phi(x)$ است زمانی‌که $0 \le y \le 1$ باشد برابر با y است و زمانی‌که $y > 1$ است ∞ می‌شود. مشابه این تعبیر برای yهای منفی نیز برقرار است. بنابراین:

$$\phi^*(y) = \begin{cases} |y| & |y| \le 1 \\ \infty & |y| > 1 \end{cases}$$

بنابراین مساله‌ی دوگان برای این تابع جریمه به‌صورت زیر نوشته می‌شود.

$$\begin{aligned} max \quad & \sum_{i=1}^{m} -|v_i| - b^T v \\ s.t \quad & A^T v = 0, |v_i| \le 1 \end{aligned} \qquad \equiv \qquad \begin{aligned} max \quad & -\|v\|_1 - b^T v \\ s.t \quad & A^T v = 0, \|v\|_\infty \le 1 \end{aligned}$$

۶- نیم‌فضایی را که در زیر تعریف شده است درنظر بگیرید. هدف یافتن نقطه‌ای داخل این نیم‌فضا است که کمترین نرم اقلیدسی را داشته باشد. این مساله را به فرم یک مساله‌ی بهینه‌سازی نوشته و با استفاده از شروط KKT حل نمایید.

$$H = \{x \in \mathbb{R}^n | a^T x + b \ge 0\} \quad where\ a \in \mathbb{R}^n, b \in \mathbb{R}$$

پاسخ:

$$\begin{aligned} min \quad & \|x\|_2^2 \\ s.t. \quad & -(a^T x + b) \le 0 \end{aligned}$$

$$\mathcal{L}(x,\lambda) = x^T x - \lambda(a^T x + b)$$

$$\nabla\mathcal{L}(x,\lambda) = 2x - \lambda a = 0 \Rightarrow x = \frac{\lambda a}{2}$$

حال با به کارگیری شرط Complementary Slackness داریم:

$$\lambda(a^T x + b) = 0$$
$$\Rightarrow \begin{cases} \lambda = 0 \Rightarrow x = 0 \Rightarrow 0 - b \le 0 \Rightarrow b \ge 0 \\ \lambda \ne 0 \Rightarrow a^T x + b = \dfrac{\lambda a^T a}{2} + b = 0 \Rightarrow \lambda = -\dfrac{2b}{\|a\|_2^2} \Rightarrow x = -\dfrac{ba}{\|a\|_2^2}, \lambda \ge 0 \Rightarrow b < 0 \end{cases}$$

۷- یک مساله‌ی بهینه‌سازی محدب بدون قیود تساوی به‌صورت زیر را درنظر بگیرید.

$$\min \ f_0(x)$$
$$s.t. \quad f_i(x) \le 0, \ \ i = 1, \dots, m$$

نشان دهید اگر x^* و λ^* شروط KKT را برآورده نمایند رابطه‌ی زیر برای تمامی x های شدنی برقرار خواهد بود.

$$\nabla f(x^*)^T(x - x^*) \ge 0$$

پاسخ:

در شروط KKT داشتیم:

$$\begin{cases} f_i(x^*) \le 0 \\ \lambda_i^* \ge 0 \\ \lambda_i^* f_i(x^*) = 0 \\ \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0 \end{cases}$$

$$0 \ge f_i(x) \ge f_i(x^*) + \nabla f_i(x^*)^T(x - x^*)$$

$$\sum_{i=1}^{m} \lambda_i f_i(x^*) + \lambda_i \nabla f_i(x^*)^T(x - x^*) \le 0$$

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x^*)^T(x - x^*) = -\nabla f_0(x^*)(x - x^*) \le 0$$

۸- مساله‌ی بهینه‌سازی محدب زیر را درنظر بگیرید.

$$\begin{aligned} &minimize && f_0(x) \\ &subject\ to && f_i(x) \le 0, \qquad i = 1, \dots, m \\ &&& Ax = b \\ &&& x^T x \le R^2 \end{aligned}$$

فرض کنید $\tilde{\phi}$ نشان‌دهنده‌ی تابع logarithmic barrier مربوط به این مساله باشد. مقدار $a > 0$ را طوری پیدا کنید که $\nabla^2\left(tf_0(x) + \tilde{\phi}(x)\right) \succcurlyeq aI$ که در آن $t > 0$ است به ازای تمامی x های شدنی[۲] برقرار باشد.

پاسخ:

فرض کند ϕ نشان دهنده‌ی تابع logarithmic barrier مساله‌ی

$$\begin{aligned} &minimize && f_0(x) \\ &subject\ to && f_i(x) \le 0, \qquad i = 1, \dots, m \\ &&& Ax = b \end{aligned}$$

باشد. قید $x^T x \le R^2$ عبارت $-\log(R^2 - x^T x)$ را به ϕ اضافه خواهد کرد. بنابراین داریم:

$$\nabla^2\left(tf_0 + \tilde{\phi}\right) = \nabla^2(tf_0 + \phi) + \frac{2}{R^2 - x^T x} I + \frac{4}{(R^2 - x^T x)^2} x x^T \succcurlyeq \nabla^2(tf_0 + \phi) + \left(\frac{2}{R^2}\right) I$$
$$\succcurlyeq \left(\frac{2}{R^2}\right) I$$

بنابراین مقدار $\frac{2}{R^2}$ می‌تواند به‌عنوان پاسخ درنظر گرفته شود.

[۲] feasible