دانشکده علوم ریاضی
دانشگاه صنعتی شریف

تحقیق در عملیات ۱

محمدهادی فروغمنداعرابی
پاییز ۱۳۹۹

آشنایی با برنامه‌ریزی محدب

جلسه بیستم

نگارنده: محمدحسین سیفی

۱ مروری بر مباحث گذشته

در چند جلسه‌ی گذشته درباره‌ی کاربردهای مختلف برنامه‌ریزی خطی صحبت کردیم و در این جلسه به آخرین کاربرد آن می‌پردازیم که در واقع تعمیم ساده‌ای از برنامه‌ریزی خطی است. البته به طور کامل برنامه‌ریزی محدب را بررسی نمی‌کنیم و با حل یک مثال به حالت خاصی از برنامه‌ریزی محدب می‌پردازیم.

۲ یک مثال: کوچکترین گوی

هدف ما پیدا کردن کوچک‌ترین گوی شامل تعدادی نقطه است.

برنامه‌ریزی محدب

هر برنامه‌ریزی خطی به صورت زیر است:

$$\begin{aligned} &\text{کمینه کن} && c'x \\ &\text{که} && Ax = b \end{aligned}$$

برنامه‌ریزی محدب یک تعمیم ساده از برنامه‌ریزی خطی است به این صورت که تابع هدف به جای یک تابع خطی، یک تابع محدب و قیود به جای یک مجموعه خطی یک مجموعه‌ی محدب است.

$$\text{کمینه کن} \quad f(x)$$
$$\text{که} \quad x \in K$$

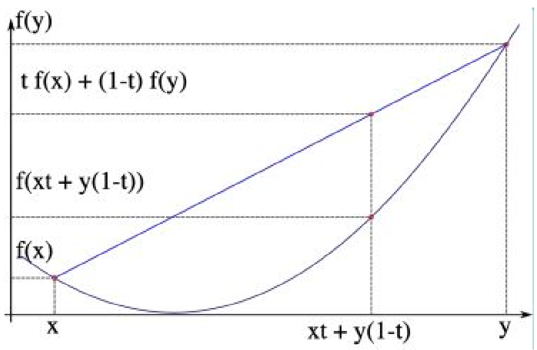
برای یادآوری مجموعه‌ی محدب و تابع محدب را تعریف می‌کنیم:

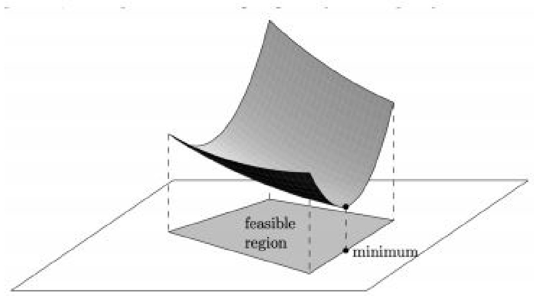
تابع محدب:

$t \in [۰, ۱]$
$f(tx + (۱ - t)y) \leq tf(x) + (۱ - t)f(y)$

مجموعه محدب:

$t \in [۰, ۱] x, y \in K$
$\implies tx + (i - t)y \in K$





شکل تعبیر هندسی تابع محدب را نشان می‌دهد و این ویژگی اصلی که مشتق در هر نقطه به نقطه کمینه اشاره میکند.
برنامه‌ریزی محدب معادل دیگری دارد که به این صورت نوشته می‌شود:

کمینه کن $f(x)$
که $g_i(x) \leq ۰, \quad i = ۱, ..., m$
$h_i(x) = ۰, \quad i = ۱, ..., p$

که در این تعریف f و g تابع محدب و h تابع خطی است.

در این جلسه ما از یک حالت میانی در برنامه‌ریزی استفاده می‌کنیم که تابع هدف محدب و قیود خطی دارد. محدویت بیشتری نسبت به برنامه‌ریزی خطی و و محدودیت کمتری نسبت به برنامه‌ریزی محدب دارد. به این صورت :

$$\text{کمینه کن} \quad f(x)$$
$$\text{که} \quad Ax = b$$
$$x \geq ۰$$

ابتدا چند قضیه درباره‌ی برنامه‌ریزی محدب بیان کرده و آن‌ها را اثبات می‌کنیم و بعد به حل یک مسئله می‌پردازیم.

قضیه

اگر f مشتق‌پذیر باشد:

$f(\lambda x + (1\lambda)y) \leq \lambda f(x) + (1)f(y)$ $\lambda \in [0, 1],$ معادل است با: $f(x) \geq f(z) + \nabla f(z)(x - z)$

شهود قضیه این است که اگر خط بین هر دو نقطه دلخواه روی محنی بالاتر از آن باشد، در هر نقطه صفحه‌ی مماس بر منحنی پایین‌تر از همه‌ی نقاص منحنی است.

اثبات:

$\leftarrow$:

$$f(x + \lambda(y - x)) \leq f(x) + \lambda(f(y) - f(x))$$
$$\implies f(y) - f(x) \geq \frac{f(x + \lambda(y - x) - f(x))}{\lambda}, \quad \forall\lambda \in (0, 1]$$
$$f(y) - f(x) \geq \nabla f^T(x)(y - x)$$

$\rightarrow$:

$$f(x) \geq f(z) + \nabla f^T(z)(x - z)$$
$$f(y) \geq f(z) + \nabla f^T(z)(y - z)$$

z را قرار می‌دهیم $z = \lambda x + (1 - \lambda y)$

λ برابر معادله اول را با $(1 - \lambda)$ برابر معادله دوم جمع می‌کنیم:

$$\lambda f(x) + (1 - \lambda f(y)) \geq f(z) + \nabla f^T(z)(\lambda x + (1 - \lambda)y - z) = f(\lambda x + (1 - \lambda)y)$$

گزاره منطقی

اگر $C \subseteq R^n$ یک مجموعه‌ی محدب و $f : R^n \Rightarrow R$ یک تابع مشتق‌پذیر باشد، بردار x^* مقدار $f(x)$ را روی C کمینه میکند اگر و تنها اگر :
$\nabla f(x^*)(x - x^*) \geq 0 \quad \forall x \in C$

اثبات:

$\leftarrow$:

اگر در قضیه‌ی قبلی به جای z قرار دهیم x^* داریم:

$$f(x) \geq f(x^*) + \nabla f(x^*)(x - x^*)$$

با توجه به اینکه داریم:

$$\nabla f(x^*)(x - x^*) \geq 0$$

به ازای هر نقطه‌ی x می‌توان گفت:

$$f(x) \geq f(x^*)$$

$\rightarrow$:

تابع x را به این صورت تعریف می‌کنیم: $x(t) := x^* + t(x - x^*) \in C, \quad t \in [0, 1]$

مشتق $f(x(t))$ را در نقطه‌ِی $t = 0$ از دو روش به دست می‌آوریم:

$$\tfrac{\partial}{\partial t} f(x(t))|_{t=0} = \lim_{t\to 0} \frac{f(x(t)) - f(x^*)}{t}$$

حد سمت راست همیشه بزرگتر مساوی صفر است چون x^* نقطه‌ی کمینه است.

و هم‌چنین داریم:

$$\tfrac{\partial}{\partial t} f(x(t))|_{t=0} = \nabla f(x^*)(x - x^*)$$

پس

$$\nabla f(x^*)(x - x^*) \geq 0$$

گزاره

Karush–Kuhn–Tucker conditions

برنامه‌ریزی محدب زیر را در نظر می‌گیریم:

$$\begin{aligned} \text{کمینه کن} \quad & f(x) \\ \text{که} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

که در آن f تابعی محدب و مشتق‌پذیر با مشتقات جزئی پیوسته است. یک جواب شدنی مثل $x^* \in R^n$ بهینه است اگر و تنها اگر بردار $\widetilde{y} \in R^m$ وجود داشته‌باشد که:

$$\begin{aligned} \nabla f(x^*)_j + \widetilde{y}^T a_j \quad &= 0 \quad if \quad x_j^* > 0 \\ \nabla f(x^*)_j + \widetilde{y}^T a_j \quad &\geq 0 \quad otherwise. \end{aligned}$$

که در آن a^j ستون j ام در ماتریس A است.

اثبات:

$\leftarrow$:

$$\begin{aligned} &(\nabla f(x^*) + \widetilde{y}^T A)x^* = 0 \\ &(\nabla f(x^*) + \widetilde{y}^T A)x \geq 0 \\ &\Longrightarrow \quad \nabla f(x^*)(x - x^*) \geq 0 \end{aligned}$$

$\rightarrow$:

x^* جواب بهینه است برای

$$\begin{aligned} \text{کمینه کن} \quad & f(x) \\ \text{که} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

داریم $\nabla f(x^*)(x - x^*) \geq 0$. پس اگر c^T تعریف کنیم $c^T = -\nabla f(x^*)$ ، به ازای هر x شدنی داریم:

$c^T(x - x^*) \leq 0$

پس x^* جواب بهینه است برای

$$\begin{aligned} \text{بیشینه کن} \quad & c^T x \\ \text{که} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

دوگان این برنامه‌ریزی به این صورت است:

$$\begin{aligned} \text{بیشینه کن} \quad & b^T x \\ \text{که} \quad & A^T y \geq c \end{aligned}$$

طبق قضیه دوگانی قوی جواب بهینه‌ای مثل $\widetilde{y}$ برای برنامه‌ریزی دوگان وجود دارد.

$$(\widetilde{y}^T A - c^T)x^* = b^T \widetilde{y} - c^T x^* = 0$$

از طرفی از $A^T y \geq c$ نتیجه می‌شود همه‌ی مولفه‌های بردار $(\widetilde{y}^T A - c^T)x^*$ نامنفی است.

$\Longrightarrow$
$$\nabla f(x^*)_j + \widetilde{y}^T a_j \quad \geq \circ$$
$$\nabla f(x^*)_j + \widetilde{y}^T a_j \quad = \circ \quad if \quad x_j^* > \circ$$

نکته‌ی قابل ذکر اینکه گرچه ما قضیه‌ی kkt را برای حالت خاصی از برنامه‌ریزی‌های محدب اثبات کردیم این قضیه برای همه‌ی برنامه‌ریزی‌های محدب صادق است.

مسئله‌ی کوچک‌ترین گوی

هدف ما پیدا کردن کوچکترین گوی در d بعد است که شامل همه‌ی n نقطه‌ی داده شده باشد.

سعی می‌کنیم یک برنامه‌ریزی محدب برای مسئله بنویسیم. $p_۱$ تا p_n نقاط ورودی هستند. متغیرهای $p \in R^b$ و $r \in R$ را به عنوان مرکز و شعاع در نظر میگیریم.

$$\text{کمینه کن} \quad r$$
$$\text{که} \quad \forall j : |p - p_j| \leq r$$

به ازای هر j قید را به این‌صورت می‌نویسیم:

$$\Sigma(p[i] - p_j[i])^۲ \leq r^۲$$
$$\Sigma p[i]^۲ + \Sigma \alpha_j^i p_j[i] + sabet \leq r^۲$$

شهود هندسی قضیه در دو بعد به این صورت است که روی هر نقطه یک سهمی‌گون به وجود می‌اید که بعد عمودی آن r است. هدف ما پیدا کرد پایین‌ترین صفحه‌ای است که همه‌ی سهمی‌گون ها در آن باهم اشتراک داشته باشند و آن نقطه‌ی اشتراک p است.

هرکدام از قیود محدب است. تابع هدف نیز محدب است. پس با یک برنامه‌ریزی محدب سر و کار داریم. اما این برناه‌ریزی از آن نوع خاصی که مد نظر ماست نیست چون تابع هدف خطی و قیود محدب اند.

ما دنبال برنامه‌ریزی با قیود خطی و تابع هدف محدبیم.

قضیه

اگر $p_۱, ..., p_n$ نقاطی در R^d و Q ماتریسی $d \times n$ باشند که ستون j ام Q درایه‌های نقطه‌ی p_j است، مسئله‌ی بهینه‌سازی زیر را در نظر بگیرید

$$\text{کمینه کن} \quad x^T Q^T Q x - \sum_{j=۱}^{n} x_j P_j^T P_j$$
$$\text{که} \quad \sum_{j=۱}^{n} x_j = ۱$$
$$x \geq \circ$$

که $x_۱, ..., x_n$ متغیر هستند. تابع هدف $f(x) := x^T Q^T Q x - \sum_{j=۱}^{n} x_j P_j^T P_j$ محدب است و داریم:

۱.برنامه‌ریزی بالا یک جواب بهینه‌ی x^* دارد.

۲. نقطه‌ی p^* وجود دارد که به ازای هر x^* به صورت $p^* = Qx^*$ تعریف می‌شود. علاوه‌براین یک گوی با مرکز p^* و با مجذور شعاع $-f(x^*)$ تنها گوی با کوچکترین شعاع شامل تمام نقاط P است.

اثبات:

شرط kkt را برای نقطه‌ی بهینه می‌نویسیم:

$$\nabla f(x^*)_j + \widetilde{y}^T a_j \quad = \circ \quad if \quad x_j^* > \circ$$
$$\nabla f(x^*)_j + \widetilde{y}^T a_j \quad \geq \circ \quad otherwise.$$

$\nabla f(x)$ را حساب می‌کنیم:

$$\nabla f(x) = ۲x^T Q^T Q - (p_۱^T p_۱, p_۲^T p_۲, ..., p_n^T p_n)$$

از تعریف $p^* = Qx^* = \sum_{j=۱}^{n} x_j^* p_j$ استفاده می‌کنیم و kkt را بازنویسی می‌کنیم:

$$
\begin{aligned}
&۲p_j^T p^* - p_j^T p_j + \mu = \circ \quad if \quad x_j^* > \circ \\
&۲p_j^T p^* - p_j^T p_j + \mu \geq \circ \quad if \quad otherwise
\end{aligned}
$$

که معادل است با

$$
\begin{aligned}
&||p_j - p^*||^۲ = \mu + {p^*}^T p^* \quad if \quad x_j^* > \circ \\
&||p_j - p^*||^۲ \leq \mu + {p^*}^T p^* \quad otherwise
\end{aligned}
$$

که $\mu + {p^*}^T p^*$ مربع شعاع است.
کافی است نشان دهیم که این عبارت تابع هدف است.

۳ ارجاع و منابع

[1] Bernard Gärtner and Jirí Matoušek. *Understanding and using linear programming.* Springer, 2007.