دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

تمرین سری دو

درس بهینه‌سازی

فرهاد دلیرانی

۹۶۱۳۱۱۲۵

dalirani@aut.ac.ir

dalirani.1373@gmail.com

فهرست

ابزارهای استفاده شده

زبان برنامه نویسی: -

محیط توسعه: -

سیستم عامل: -

مسائل تشریحی _ سوال 1

$$f(x_k) + (1-c)\alpha_k \nabla f_k^T p_k < f(x_k + \alpha_k p_k) \leq f(x_k) + c\alpha_k \nabla f_k^T p_k$$

این نامساوی در کتاب Nocedal ویرایش دو آمده است. این نامساوی با عنوان The Goldstein Conditions شناخته می‌شود.

اگر به نامساوی سمت راست $[f(x_k+\alpha_k p_k) \leq f(x_k)+c\alpha_k \nabla f_k^T p_k]$ نگاه کنیم ، متوجه می‌شویم نامساوی سمت راست یکی از شرایط ولف است که نشان می‌دهد مقدار کاهش در f متناسب با اندازه‌ی گام و $\phi'(0) = \nabla f_k^T p_k$ است و مقدار کاهش f را تضمین می‌کند. (sufficient decrease)

نامساوی سمت چپ $[f(x_k) + (1-c)\alpha_k \nabla f_k^T p_k \leq f(x_k+\alpha_k p_k)]$ بیان می‌کند که میزان α نباید از حد مشخصی کمتر شود. در واقع نامساوی سمت چپ step length را از پایین کنترل می‌کند. این موضوع را در تصویر زیر مشاهده می‌کنید و مانع کوچک شدن زیاد α می‌شود.

یک تفاوت دیگر این شرایط با شرایط ولف این است که این شرایط فقط یک constant دارد ولی شرایط ولف دو constant دارد.

از بدی های این روش این است ممکن است minimizer های تابع $\phi(\alpha) = f(x_k+\alpha p_k)$, $\alpha > 0$ ممکن است در شرایط Goldstein صدق نکنند.

از این روش معمولا برای متدهای نیوتنی استفاده می‌شود و معمولا برای روش‌های quasi Newton مناسب نیست.

به طور خلاصه ، شرایط Goldstein اطمینان حاصل می‌کند که α (step length) منجر به sufficient decrease می‌شود ولی برای α یک حد پایین مشخص می‌کند و مانع بسیار کوچک شدن آن می‌شود.

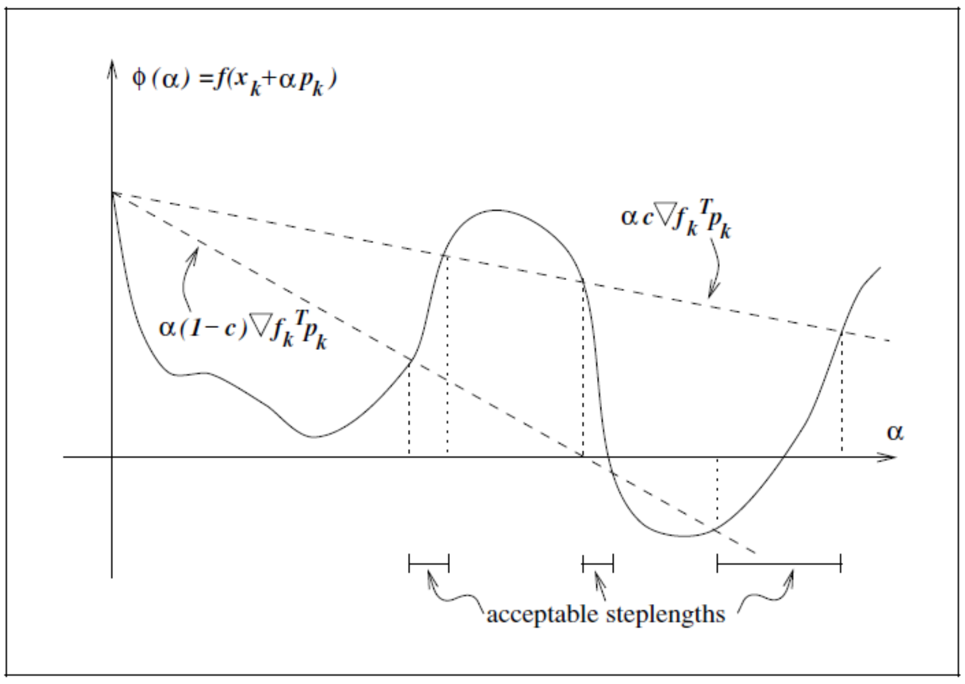

Figure 3.6 The Goldstein conditions.

تصویر از صفحه ۳۶ کتاب Numerical Optimization نوشته شده توسط J.Nocedal ویراش دو برداشته شده است.

مسائل تشریحی - سوال 2

$$P_k = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad x_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$f(x) = (x_1 + x_2^2)^2$$

$$\Longrightarrow \nabla f(x) = \begin{bmatrix} 2(x_1 + x_2^2) \\ 2(x_1 + x_2^2)(2x_2) \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2^2 \\ 4x_1x_2 + 2x_2^3 \end{bmatrix}$$

$$\Longrightarrow \nabla f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \Longrightarrow -\nabla f_k = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

زاویه بین P_k و $-\nabla f_k$ $\Longrightarrow \cos\theta = \dfrac{-\nabla f_k^T P_k}{\|\nabla f_k\| \cdot \|P_k\|} = \dfrac{2}{2\sqrt{2}} = \dfrac{\sqrt{2}}{2}$

$\Longrightarrow \theta = \dfrac{\pi}{4} \longrightarrow$ زاویه‌ی P_k و $-\nabla f_k$ کمتر از نود درجه است

به همین دلیل P_k یک جهت Descent است.

برای بدست آوردن مقدار α :

$$f(x_k + \alpha P_k) = f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = f\left(\begin{bmatrix} 1-\alpha \\ \alpha \end{bmatrix}\right) = \left((1-\alpha) + \alpha^2\right)^2$$

برای به دست آوردن α ، مشتق می‌گیریم و سپس مساوی صفر قرار می‌دهیم.

$2(1-\alpha+\alpha^2)(-1+2\alpha) = 0 \longrightarrow \alpha^2 - \alpha + 1 = 0 \rightarrow \Delta = 1-4 = -3 < 0$ جواب ندارد

$\longrightarrow \boxed{\alpha = \dfrac{1}{2}}$ ✓

مسائل تشریحی ـ سوال ۳

(۲).

$$\min \|Ax-b\|_1 \qquad A_{m\times n} \quad x_{n\times 1} \quad b_{m\times 1}$$

$$\text{subject to} \quad \|x\|_\infty \leqslant 1$$

$$Ax-b = \begin{bmatrix} \text{---}\ a_1^T\ \text{---} \\ \text{---}\ a_2^T\ \text{---} \\ \vdots \\ \text{---}\ a_m^T\ \text{---} \end{bmatrix} x - b = \begin{bmatrix} a_1^T x - b_1 \\ a_2^T x - b_2 \\ \vdots \\ a_m^T x - b_m \end{bmatrix}$$

$$\rightarrow \|Ax-b\|_1 = |a_1^T x - b_1| + |a_2^T x - b_2| + \dots + |a_m^T x - b_m|$$

$$\begin{aligned} &\min \|Ax-b\|_1 \\ &\text{subject to} \quad \|x\|_\infty \leqslant 1 \end{aligned} \Longrightarrow \begin{aligned} &\min |a_1^T x - b_1| + |a_2^T x - b_2| + \dots + |a_m^T x - b_m| \\ &\text{subject to} \quad \|x\|_\infty \leqslant 1 \end{aligned}$$

به فرم Epigraph می‌نویسیم

$$\longrightarrow \begin{aligned} &\min \ t \\ &\text{subject to} \quad |a_1^T x - b_1| + |a_2^T x - b_2| + \dots + |a_m^T x - b_m| \leqslant t \\ &\qquad\qquad\quad \|x\|_\infty \leqslant 1 \end{aligned}$$

$$\boxed{t_1 + t_2 + t_3 + \dots + t_m = t}$$

$$\longrightarrow \begin{aligned} &\min \ t_1 + t_2 + t_3 + \dots t_m \\ &\text{subject to} \quad |a_1^T x - b_1| \leqslant t_1 \\ &\qquad\qquad\quad |a_2^T x - b_2| \leqslant t_2 \\ &\qquad\qquad\quad \vdots \\ &\qquad\qquad\quad |a_m^T x - b_m| \leqslant t_m \\ &\qquad\qquad\quad \|x\|_\infty \leqslant 1 \end{aligned}$$

ادامه

$$\|x\|_\infty = \max\left(|x_1|, |x_2|, \dots |x_n|\right)$$

$$\|x\|_\infty \leq 1 \rightarrow \max\left(|x_1|, |x_2|, \dots, |x_n|\right) \leq 1$$

$$\rightarrow \quad |x_1| \leq 1 \,, \ |x_2| \leq 1 \,, \ \dots \,, \ |x_n| \leq 1$$

$$\min \ t_1 + t_2 + \dots + t_m$$

$$\text{subject to} \quad \begin{aligned} -t_1 &\leq a_1^T x - b_1 \leq t_1 \\ -t_2 &\leq a_2^T x - b_2 \leq t_2 \\ &\ \vdots \\ -t_m &\leq a_m^T x - b_m \leq t_m \\ -1 &\leq x_1 \leq 1 \\ -1 &\leq x_2 \leq 1 \\ &\ \vdots \\ -1 &\leq x_n \leq 1 \end{aligned}$$

$$T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_m \end{bmatrix} \qquad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\min \ \mathbf{1}^T T$$

$$\text{subject to} \quad -T \preceq Ax - b \preceq T$$

$$-\mathbf{1} \preceq x \preceq \mathbf{1}$$

affine

affine

affine

به فرم LP درآمد.

مسائل تشریحی - سوال ۳

(ب)

$$\min \|Ax-b\|_\infty \qquad A_{m\times n}$$

$$A = \begin{bmatrix} — a_1^T — \\ — a_2^T — \\ \vdots \\ — a_m^T — \end{bmatrix} \qquad a_i^T \text{ are rows of } A$$

$$Ax-b = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} x - b = \begin{bmatrix} a_1^T x - b_1 \\ a_2^T x - b_2 \\ \vdots \\ a_m^T x - b_m \end{bmatrix}$$

$$\|Ax-b\| = \max\left(|a_1^T x - b_1|, |a_2^T x - b_2|, |a_3^T x - b_3|, \dots, |a_m^T x - b_m|\right)$$

$$\min \|Ax-b\|_\infty \xrightarrow{\text{Epi graph فرم}} \begin{array}{ll} \min & t \\ \text{subject to} & \|Ax-b\|_\infty \leq t \end{array}$$

$$\longrightarrow \begin{array}{ll} \min & t \\ \text{subject to} & \max(|a_1^T x - b_1|, |a_2^T x - b_2|, \dots, |a_m^T x - b_m|) \leq t \end{array}$$

$$\boxed{\begin{array}{c} P_1 \leq P_2 \leq P_3 \leq \cdots \leq P_k \\ P_k \leq q \end{array} \Longrightarrow \begin{array}{c} P_1 \leq q \\ P_2 \leq q \\ \vdots \\ P_k \leq q \end{array}} \Longrightarrow \begin{array}{ll} \min & t \\ \text{subject to} & |a_1^T x - b_1| \leq t \\ & |a_2^T x - b_2| \leq t \\ & \vdots \\ & |a_m^T x - b_m| \leq t \end{array}$$

$$\longrightarrow \begin{array}{ll} \min & t \\ \text{subject to} & -t \leq a_1^T x - b_1 \leq t \\ & -t \leq a_2^T x - b_2 \leq t \\ & \vdots \\ & -t \leq a_m^T x - b_m \leq t \end{array} \xRightarrow{\boxed{\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}} \begin{array}{ll} \min & t \\ \text{subject to} & -t\mathbf{1} \preceq Ax-b \preceq t\mathbf{1} \end{array}$$

$$\Longrightarrow \begin{array}{ll} \min & t & \longrightarrow \text{affine} \\ \text{subject to} & Ax-b \preceq t\mathbf{1} & \longrightarrow \text{affine} \\ & -t\mathbf{1} \preceq Ax-b & \longrightarrow \text{affine} \end{array} \longrightarrow \text{LP}$$

به فرم LP در آمد

تمرین۴

مسائل تشریحی - سوال 4

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \qquad \text{Show} \quad p^* = \begin{cases} c^T A^{-1} b & A^{-T}c \preceq 0 \\ -\infty & \text{otherwise} \end{cases} \qquad (4.9)$$

$$t = Ax \xrightarrow{A \text{ is nonsingular}} x = A^{-1}t$$

$$\Rightarrow \begin{array}{ll} \text{minimize} & c^T A^{-1} t \\ \text{subject to} & t \preceq b \end{array} \Rightarrow \begin{array}{ll} \text{minimize} & (A^{-T}c)^T t \\ \text{subject to} & t \preceq b \end{array}$$

اگر $A^{-T}c \succ 0$ باشد، هرچه المنت های وکتور t را کوچک تر انتخاب کنیم مقدار $(A^{-T}c)^T t$ مقدار کوچک تری شود. از آن جایی که t حد پایین ندارد می توان t را به صورت $\begin{bmatrix} -\infty \\ -\infty \\ \vdots \\ -\infty \end{bmatrix}$ انتخاب کرد، در نتیجه در این صورت LP به صورت unbounded below است.

اگر $A^{-T}c \preceq 0$ باشد، هرچه t کوچک تری انتخاب کنیم مقدار $(A^{-T}c)^T t$ بزرگ تر می شود و در $t = \begin{bmatrix} -\infty \\ -\infty \\ \vdots \\ -\infty \end{bmatrix}$ مقدار $(A^{-T}c)^T t$ به سمت مثبت بی نهایت می رود.

در این صورت کمترین مقدار $(A^{-T}c)^T t$ زمانی رخ می دهد که بزرگ ترین مقدار t انتخاب شود که این مقدار برابر با $t = b$ است در نتیجه:

$$p^* = \begin{cases} (A^{-T}c)^T b & A^{-T}c \preceq 0 \\ -\infty & \text{Else} \end{cases}$$

or

$$p^* = \begin{cases} c^T A^{-1} b & A^{-T}c \preceq 0 \\ -\infty & \text{Else} \end{cases}$$

مسائل تشریحی - سوال 4

(4.12)

جریان خارجی به هر گره به صورت b_i است که اگر مثبت باشد نشان می دهد b_i واحد از خارج وارد می شود. اگر منفی باشد نشان می دهد b_i واحد به خارج داده می شود.

جریان های ورودی به یک گره i به صورت $\sum_{j=1}^{n} x_{ji}$ است.

جریان های خروجی از گره i به صورت $\sum_{j=1}^{n} x_{ij}$ است.

از آنجایی که مجموع جریان هایی که به گره وارد و خارج شده اند برابر است، می توان معادله ی زیر را نوشت:

$$b_i + \sum_{j=1}^{n} x_{ji} - \sum_{j=1}^{n} x_{ij} = 0$$

هر لبه جهت دار x_{ij} می تواند مقادیر زیر را داشته باشد

$$l_{ij} \leqslant x_{ij} \leqslant u_{ij}$$

در نتیجه مسئله بهینه سازی به شکل زیر در می آید:

$$\text{minimize} \quad C = \sum_{i,j=1}^{n} c_{ij} x_{ij}$$

$$\text{subject to} \quad b_i + \sum_{j=1}^{n} x_{ji} - \sum_{j=1}^{n} x_{ij} = 0 \qquad (i = 1, \dots, n)$$

$$l_{ij} \leqslant x_{ij} \leqslant u_{ij}$$