بهینه سازی
مبانی بهینه سازی نامقید ذیل

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
  - مزدوج گرادیان
  - نیوتن
  - شبه‌نیوتن

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
    - $\boldsymbol{p}_i = -\nabla f(\boldsymbol{x}_i)$
    - $\boldsymbol{p}_i = -\mathrm{I}\nabla f(\boldsymbol{x}_i)$
    - I ماتریس همانی
  - مزدوج گرادیان
  - نیوتن
  - شبه‌نیوتن

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
  - مزدوج گرادیان
    - $\boldsymbol{p}_i = -\nabla f(\boldsymbol{x}_i) + \boldsymbol{\beta}_i \boldsymbol{p}_{i-1}$
      - ریوز فلچر
      - پولاک-ریبریه
  - نیوتن
  - شبه‌نیوتن

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
  - مزدوج گرادیان
  - نیوتن
    - $\boldsymbol{p}_i = -H^{-1} \nabla f(\boldsymbol{x}_i)$
      - $H = \nabla^2 f(\boldsymbol{x}_i)$
    - نیوتن تغییریافته
      - $\boldsymbol{p}_i = -(H + \nu I)^{-1} \nabla f(\boldsymbol{x}_i)$
  - شبه‌نیوتن

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
  - مزدوج گرادیان
  - نیوتن
  - شبه‌نیوتن
    - $\boldsymbol{p}_i = -B_i^{-1} \nabla f(\boldsymbol{x}_i)$
    - B_i تقریب ماتریس هسی
    - با استفاده از ماتریس رتبه پایین
      - مر-۱
      - دفپ
      - بفگش

روش‌های جستجو خط

- یافتن مسیر جستجو $\boldsymbol{p}_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$x_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$

- α_i
  - طول قدم
  - سرعت یادگیری
- $\boldsymbol{p}_i$
  - تندترین نزول
  - مزدوج گرادیان
  - نیوتن
  - شبه‌نیوتن
- همه ماتریس‌ها متقارن و ناتکین
- مثبت معین

$$\boldsymbol{p}_i^T \nabla f(\boldsymbol{x}_i) = -\nabla f(\boldsymbol{x}_i)^T B_i^{-1} \nabla f(\boldsymbol{x}_i) < 0$$

پس $\boldsymbol{p}_i$ مسیری نزولی

همگرائی روش‌های جستجو خط

الگوریتم‌های عددی تکراری (مرحله‌به‌مرحله)

بهینه‌سازی مناسب بسته به انتخاب طول قدم مناسب و جهت مناسب

همگرائی به جهت می‌پردازد

از زاویه دیگر

- همگرائی به معنی رسیدن به نقطه کمینه

همگرائی سراسری:

- الگوریتم با شروع از هر فاصلهٔ دوری از کمینه، به کمینه برسد

- همگرایی سراسری روش‌های جستجو خطی $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$
  - به نقطه مانا
  - مگر استفاده از اطلاعات هسی

مرتبه همگرائی

کوچکترین کران بالای اعداد نامنفی p صادق در

$$0 \leqslant \overline{\lim_{k\to\infty}} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} < \infty$$

ملاک بررسی در حین میل به بی‌نهایت

- ذیل مقادیر و مراحل

کاهش فاصله از حد به اندازه توان p

مرتبه همگرائی

فرض وجود دنباله و حد زیر

$$\beta = \lim_{k\to\infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p}$$

آن‌گاه

$$|r_{k+1} - r^*| = \beta |r_k - r^*|^p$$

مرتبه همگرائی- مثال

$$\beta = \lim_{k\to\infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p}$$

$0 < a < 1$

الف) $r_k = a^k$

- همگرا به صفر با مرتبهٔ یک
- $\frac{r_{k+1}}{r_k} = a$

ب) $r_k = a^{(2^k)}$

- همگرا به صفر با مرتبه دو
- $\frac{r_{k+1}}{{r_k}^2} = a$

همگرائی خطی

$$\lim_{k\to\infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|} = \beta < 1$$

همگرا به حد با نسبت (سرعت) همگرایی β

سرعت $c\beta^k$

- همگرائی هندسی!

مثال- مقایسه دو الگوریتم خطی

- مقایسه بر حسب نسبت‌های همگرائی متناظر آن‌ها
- نسبت کوچکتر نمایشگر سرعت بیشتر

$\beta = 0$ همگرائی زبرخطی

همگرائی روش‌های جستجو خط

تندترین نزول همیشه در راستای کاهش گرادیان پس همگرا به نقطه مانا

روش نیوتن-محور

- در صورتی که ماتریس هسی یا تقریب هسی
  - دارای عدد شرط حددار
  - مثبت معین
- رعایت شروط وولف

روش گرادیان مزدوج

سرعت همگرائی

همگرائی آسان به دنبال مسیرهای غیرمتعامد با گرادیان

- ولی کافی نیست
- گاهی اوقات موجب اشتباه

بسیار کند

- تندترین نزول

روش نیوتن رسیدن به جواب وقتی در همسایگی باشد

- ولی در دور دست لزوما به جواب نخواهد رسید

سرعت همگرایی گرادیان نزولی

محاسبهٔ طول قدم کمنیه‌ساز تابع $f(x_k - \alpha \nabla f_k)$

$$f(x_k - \alpha \nabla f_k) = \frac{1}{2}(x_k - \alpha \nabla f_k)^T Q (x_k - \alpha \nabla f_k) - b^T (x_k - \alpha \nabla f_k)$$

با مشتق‌گیری نسبت به طول قدم و برابر صفر قرار دادن

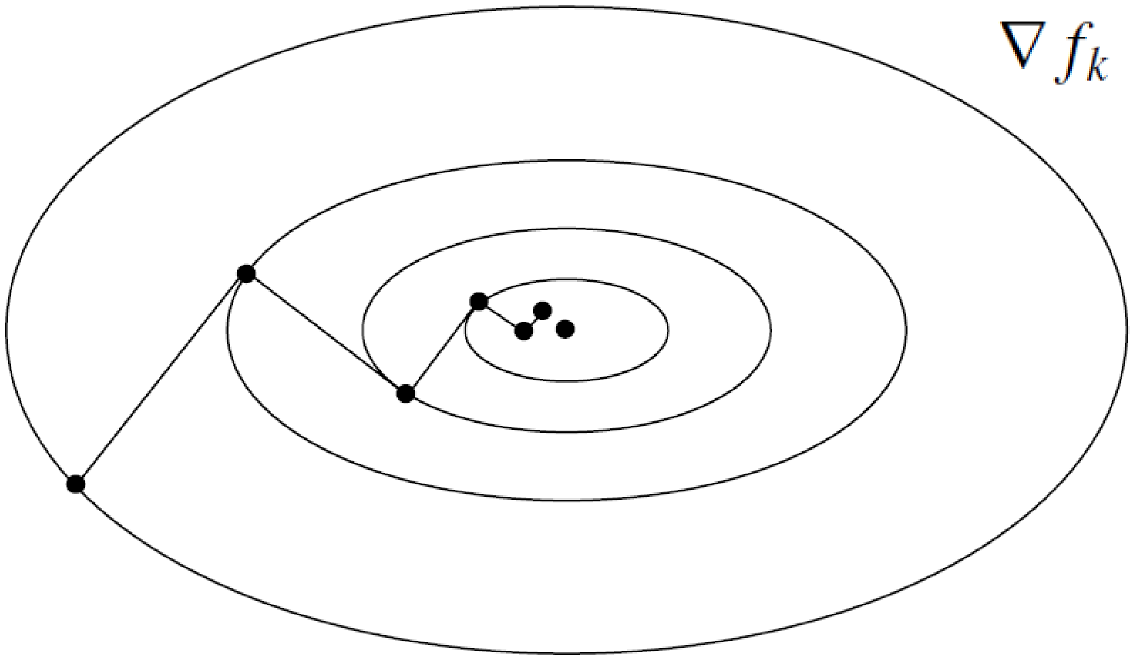
$$\alpha_k = \frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k}$$

سرعت همگرایی گرادیان نزولی

با استفاده از مقدار دقیق کمینه‌ساز طول قدم

$$x_{k+1} = x_k - \left(\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k} \right) \nabla f_k$$

$$\nabla f_k = Q x_k - b$$



امکان نوشتن صورت بسته x_{k+1} بر اساس x_k

بیضی‌وارهای **n**-بعدی

سرعت همگرایی گرادیان نزولی

$$\|x\|_Q^2 = x^T Q x$$

$$Q x^* = b$$

$$\frac{1}{2}\|x - x^*\|_Q^2 = f(x) - f(x^*)$$

سرعت همگرایی گرادیان نزولی

$$\nabla f_k = Q(x_k - x^*)$$

$$x_{k+1} = x_k - \left(\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k}\right) \nabla f_k$$

$$\|x_{k+1} - x^*\|_Q^2 = \left\{1 - \frac{\left(\nabla f_k^T \nabla f_k\right)^2}{\left(\nabla f_k^T Q \nabla f_k\right)\left(\nabla f_k^T Q^{-1} \nabla f_k\right)}\right\} \|x_k - x^*\|_Q^2$$

مقدار دقیق کاهش **f** در هر مرحله تکرار الگوریتم

- سختی تفسیر

سرعت همگرایی گرادیان نزولی

استفاده از عدد شرط

قضیه

روش تندترین نزول با جستجو خط دقیق روی تابع درجه دو محدب

$$\|x_{k+1} - x^*\|_Q^2 \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \|x_k - x^*\|_Q^2$$

بردارهای ویژه Q

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

سرعت همگرایی گرادیان نزولی

$$\|x_{k+1} - x^*\|_Q^2 \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \|x_k - x^*\|_Q^2$$

نشانگر سرعت خطی همگرایی

مورد خاص

- بردارویژه‌ها با هم برابر
  - همگرایی با یک تکرار
  - سطح ترازها دایره‌ای

سرعت همگرایی گرادیان نزولی

$$\|x_{k+1} - x^*\|_Q^2 \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \|x_k - x^*\|_Q^2$$

عدد شرط

$$\kappa(Q) = \lambda_n / \lambda_1$$

- افزایش عدد شرط
  - کشیده‌شدن سطح‌ترازها
  - افزایش زیگزاگ‌ها
  - نزول سرعت همگرایی
  - عدد شرط بزرگ برابر با ماتریس بدتعریف

سرعت همگرایی گرادیان نزولی

$$r \in \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}, 1\right)$$

$$f(x_{k+1}) - f(x^*) \leq r^2[f(x_k) - f(x^*)]$$

همگرایی روش نیوتن

همگرایی محلی

مجانبی درجه دو

با فرض

- $H(\boldsymbol{x}^*) \succ 0$
- پیوسته و هموار

انواع روش‌های محاسبه خطا

$$\delta_k = \|f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)\| \text{ یا } \delta_k = \|\boldsymbol{x}_k - \boldsymbol{x}^*\|$$

$$\delta_{k+۱} \leq c\delta_k^۲$$

مثال $\delta_k = ۱۰^{-۲}$

- $\delta_{k+۱} = ۱۰^{-۴}$ و $\delta_{k+۲} = ۱۰^{-۸}$

همگرایی روش نیوتن

$$\|x_k + p_k^{\mathrm{N}} - x^*\| \leq \tilde{L}\|x_k - x^*\|^2$$

همگرایی روش شبه‌نیوتن

$$\|x_k + p_k - x^*\| \leq o(\|x_k - x^*\|)$$

سرعت همگرائی گرادیان مزدوج

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\|_Q^2 \leq \left(\frac{\lambda_{n-k} - \lambda_1}{\lambda_{n-k} + \lambda_1}\right)^2 \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_Q^2$$

سرعت همگرائی گرادیان مزدوج

$$\|\boldsymbol{x}_{m+1} - \boldsymbol{x}^*\|_Q \approx \epsilon \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_Q$$

سرعت همگرائی گرادیان مزدوج

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_Q \leq 2\left(\frac{\sqrt{\kappa(Q)} - 1}{\sqrt{\kappa(Q)} + 1}\right)^2 \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_Q$$

سرعت همگرائی گرادیان مزدوج

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\| \leq \frac{\sqrt{\lambda_n} - \sqrt{\lambda_1}}{\sqrt{\lambda_n} + \sqrt{\lambda_1}} \|\boldsymbol{x}_k - \boldsymbol{x}^*\|$$

در صورت غلبه λ_n بر λ_1

$$\frac{\sqrt{\lambda_n} - \sqrt{\lambda_1}}{\sqrt{\lambda_n} + \sqrt{\lambda_1}} = 1 - 2\sqrt{\frac{\lambda_1}{\lambda_n}}$$

منابع

[نازه‌دل]

[لوئنبرگر]