



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  $p_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$x_{i+1} = x_i + \alpha_i p_i$$

$\alpha_i$  ▪

- طول قدم
- سرعت یادگیری

$p_i$  ▪

- تندترین نزول
- مزدوج گرادیان
- نیوتن
- شبه‌نیوتن

# روش‌های جستجو خط

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

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

▪  $\alpha_i$

- طول قدم
- سرعت یادگیری

▪  $\mathbf{p}_i$

- تندترین نزول

$$\mathbf{p}_i = -\nabla f(\mathbf{x}_i)$$

$$\mathbf{p}_i = -I \nabla f(\mathbf{x}_i)$$

- $I$  ماتریس همانی

- مزدوج گرادیان

- نیوتن

- شبه نیوتن

# روش‌های جستجو خط

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

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

▪  $\alpha_i$

- طول قدم
- سرعت یادگیری

▪  $\mathbf{p}_i$

- تندترین نزول
- مزدوج گرادیان

$$\mathbf{p}_i = -\nabla f(\mathbf{x}_i) + \beta_i \mathbf{p}_{i-1}$$

- ریوز فلچر
- پولاک-ریبری

▪ نیوتن

▪ شبه‌نیوتن

# روش‌های جستجو خط

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

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

▪  $\alpha_i$

- طول قدم
- سرعت یادگیری

▪  $\mathbf{p}_i$

- تندترین نزول
- مزدوج گرادیان
- نیوتن

$$\mathbf{p}_i = -H^{-1} \nabla f(\mathbf{x}_i) \quad \square$$

$$H = \nabla^2 f(\mathbf{x}_i) \quad \square$$

- نیوتن تغییر یافته

$$\mathbf{p}_i = -(H + \nu I)^{-1} \nabla f(\mathbf{x}_i) \quad \square$$

- شبه نیوتن

# روش‌های جستجو خط

- یافتن مسیر جستجو  $p_i$
- تصمیم بر اندازه قدم در راستای مسیر جستجو

$$x_{i+1} = x_i + \alpha_i p_i$$

$\alpha_i$  ▪

- طول قدم
- سرعت یادگیری

$p_i$  ▪

- تندترین نزول
- مزدوج گرادیان
- نیوتن
- شبه نیوتن
- $p_i = -B_i^{-1} \nabla f(x_i)$
- $B_i$  تقریب ماتریس هسی
- با استفاده از ماتریس رتبه پایین
- م-م ۱
- دفپ
- ب فگش

# روش‌های جستجو خط

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

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

▪  $\alpha_i$

- طول قدم
- سرعت یادگیری

▪  $\mathbf{p}_i$

- تندترین نزول
- مزدوج گرادیان
- نیوتن
- شبه نیوتن
- همه ماتریس‌ها متقارن و ناتکین
- مثبت معین

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

پس  $\mathbf{p}_i$  مسیری نزولی

# همگرایی روش های جستجو خط

الگوریتم های عددی تکراری (مرحله به مرحله)

بهینه سازی مناسب بسته به انتخاب طول قدم مناسب و جهت مناسب

همگرایی به جهت می پردازد

از زاویه دیگر

▪ همگرایی به معنی رسیدن به نقطه کمینه

همگرایی سراسری:

▪ الگوریتم با شروع از هر فاصله دوری از کمینه، به کمینه برسد

▪ همگرایی سراسری روش های جستجو خطی  $x_{i+1} = x_i + \alpha_i p_i$

▪ به نقطه مانا

▪ مگر استفاده از اطلاعات هسی



# مرتبۀ همگرایی

کوچکترین کران بالای اعداد نامنفی  $p$  صادق در

$$0 \leq \overline{\lim}_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} < \infty$$

ملاک بررسی در حین میل به بی‌نهایت

▪ ذیل مقادیر و مراحل

کاهش فاصله از حد به اندازه توان  $p$

# مرتبۀ همگرایی

فرض وجود دنباله و حد زیر

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

آن‌گاه

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

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

## مرتبہ همگرائی - مثال

$$0 < a < 1$$

$$r_k = a^k \text{ (الف)}$$

▪ همگرا به صفر با مرتبہ یک

$$\frac{r_{k+1}}{r_k} = a \text{ ▪}$$

$$r_k = a^{(2^k)} \text{ (ب)}$$

▪ همگرا به صفر با مرتبہ دو

$$\frac{r_{k+1}}{r_k^2} = a \text{ ▪}$$

# همگرایی خطی

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

همگرا به حد با نسبت (سرعت) همگرایی  $\beta$

سرعت  $c\beta^k$

▪ همگرایی هندسی!

مثال - مقایسه دو الگوریتم خطی

▪ مقایسه بر حسب نسبت‌های همگرایی متناظر آنها

▪ نسبت کوچکتر نمایشگر سرعت بیشتر

$\beta = 0$  همگرایی زبرخطی

# همگرایی روش های 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 همگرایی

همگرایی آسان به دنبال مسیرهای غیرمتعامل 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 لزوماً به جواب نخواهد رسید

# سرعت همگرایی گرادیان نزولی

محاسبه طول قدم کمینه‌ساز تابع  $f(x_k - \alpha \nabla f_k)$

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

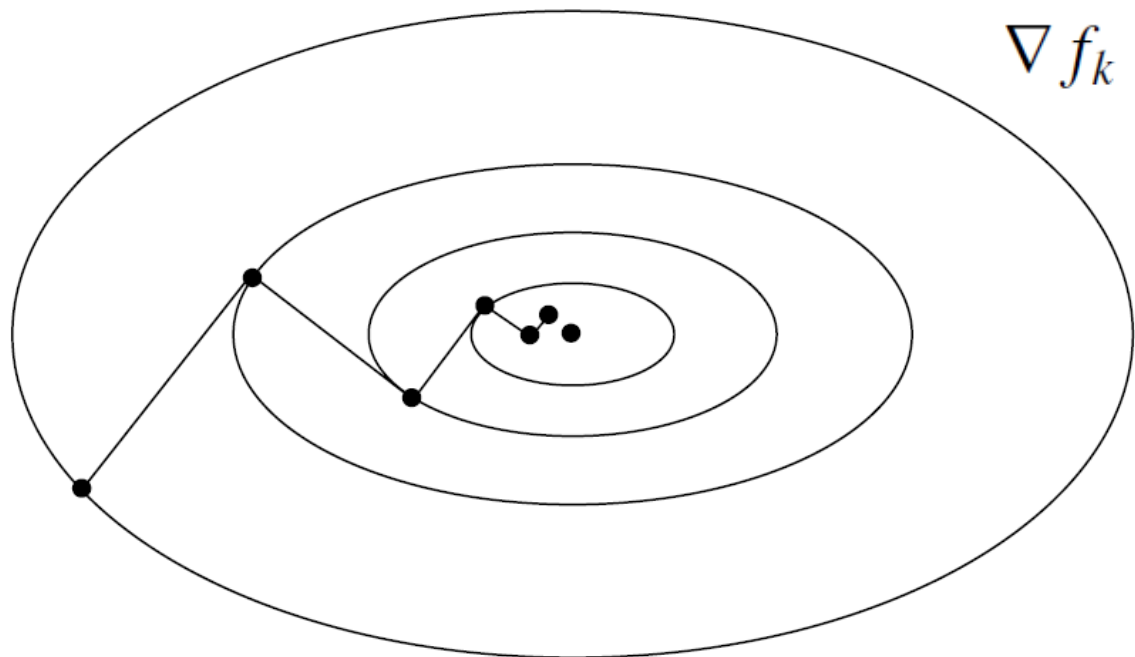
با مشتق‌گیری نسبت به طول قدم و برابر صفر قرار دادن

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

# سرعت همگرایی گرادینان نزولی

با استفاده از مقدار دقیق کمینه‌ساز طول قدم

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



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

بیضی‌وارهای n-بعدی



# سرعت همگرایی گرادیان نزولی

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

$$Qx^* = b$$

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

# سرعت همگرایی گرادیان نزولی

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

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

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

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  $Q$

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

# 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

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

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

عدد شرط

- افزایش عدد شرط
- کشیده شدن سطح ترازها
- 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

$$H(\mathbf{x}^*) > 0 \quad \bullet$$

▪ پیوسته و هموار

انواع روش‌های محاسبه خطا

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

$$\delta_{k+1} \leq c\delta_k^2$$

مثال  $\delta_k = 10^{-2}$

$$\delta_{k+1} = 10^{-4} \text{ و } \delta_{k+2} = 10^{-6} \quad \bullet$$

# همگرایی روش نیوتن

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



# همگرایی روش شبه نیوتن

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

# سرعت همگرایی گرادیان مزدوج

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

# سرعت همگرایی گرادیان مزدوج

$$\|x_{m+1} - x^*\|_Q \approx \epsilon \|x_0 - x^*\|_Q$$

# سرعت همگرایی گرادیان مزدوج

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

# سرعت همگرایی گرادیان مزدوج

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

در صورت غلبه  $\lambda_n$  بر  $\lambda_1$

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

# منابع

[نازهدل]

[لوئیسرگر]