دانشگاه صنعتی اصفهان
دانشکده مهندسی حمل و نقل

تحلیل سیستم های حمل و نقل

مروری بر برخی الگوریتم های بهینه سازی

مسائل چندمتغیره بدون محدودیت

مدرس: محمد تمنایی

بهار ۱۳۹۶

فهرست:

- ✓ الگوریتمهای کمینه سازی تک متغیره
- ✓ الگوریتمهای کمینه سازی چندمتغیره بدون محدودیت
- ✓ الگوریتمهای کمینه سازی چندمتغیره با محدودیت
- ✓ روش ترکیب محدب

مسائل چندمتغیره

از هر نقطه X^n به نقطه دیگری چون X^{n+1} حرکت می کنیم، بنحوی که: $Z(x^n) < Z\left(x^{n+1}\right)$ و:

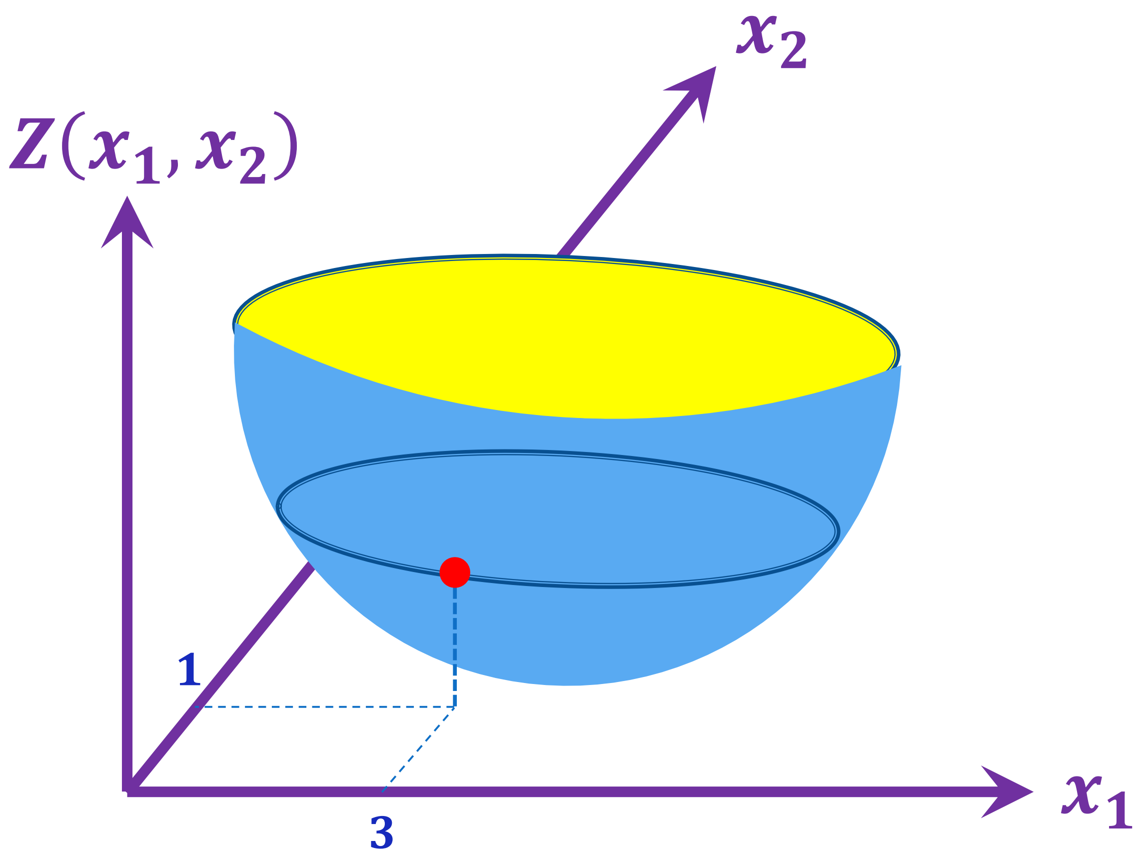
$$X^{n+1} = X^n + \alpha_n \overrightarrow{d^n}$$

$\overrightarrow{d^n}$ بردارِ جهت حرکت در تکرار n ام

α_n اسکالر بزرگی گام حرکت در تکرار n ام

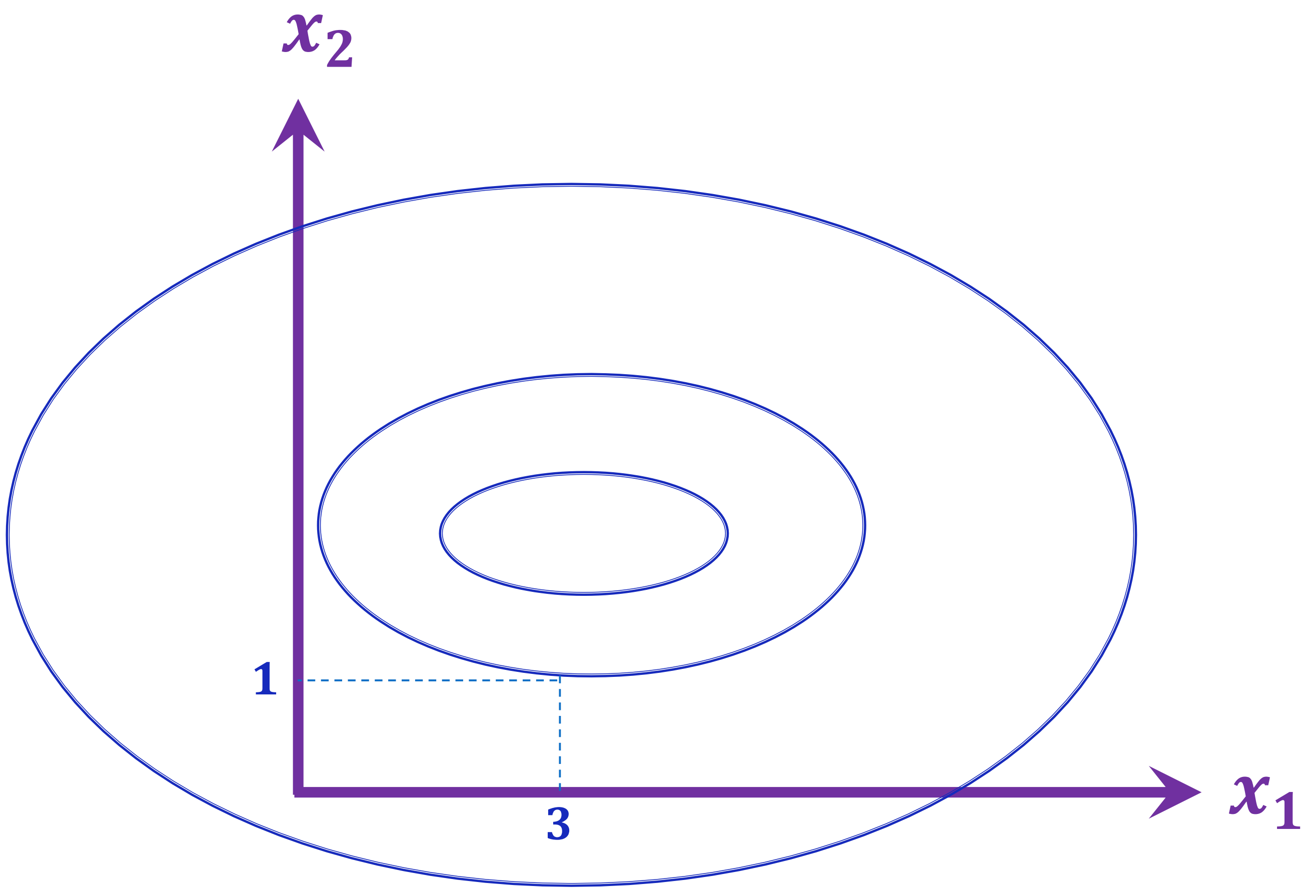
مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$



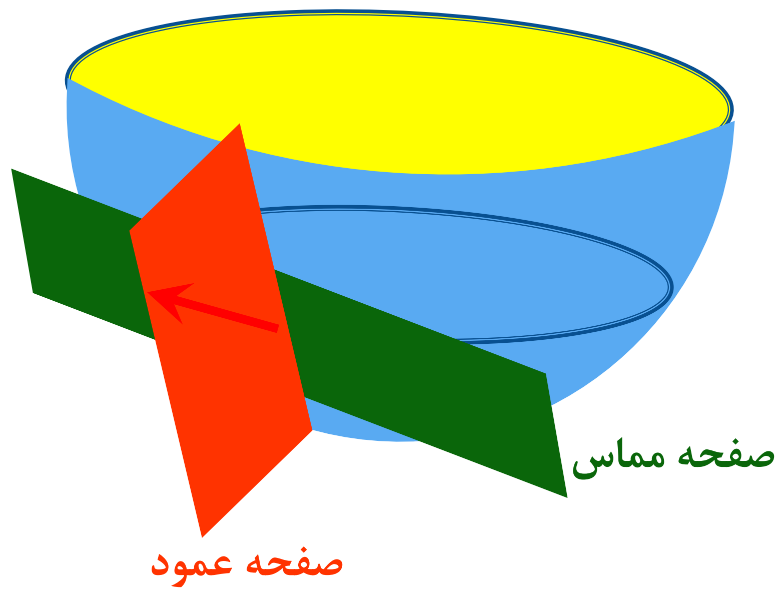
مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$



قضیه:

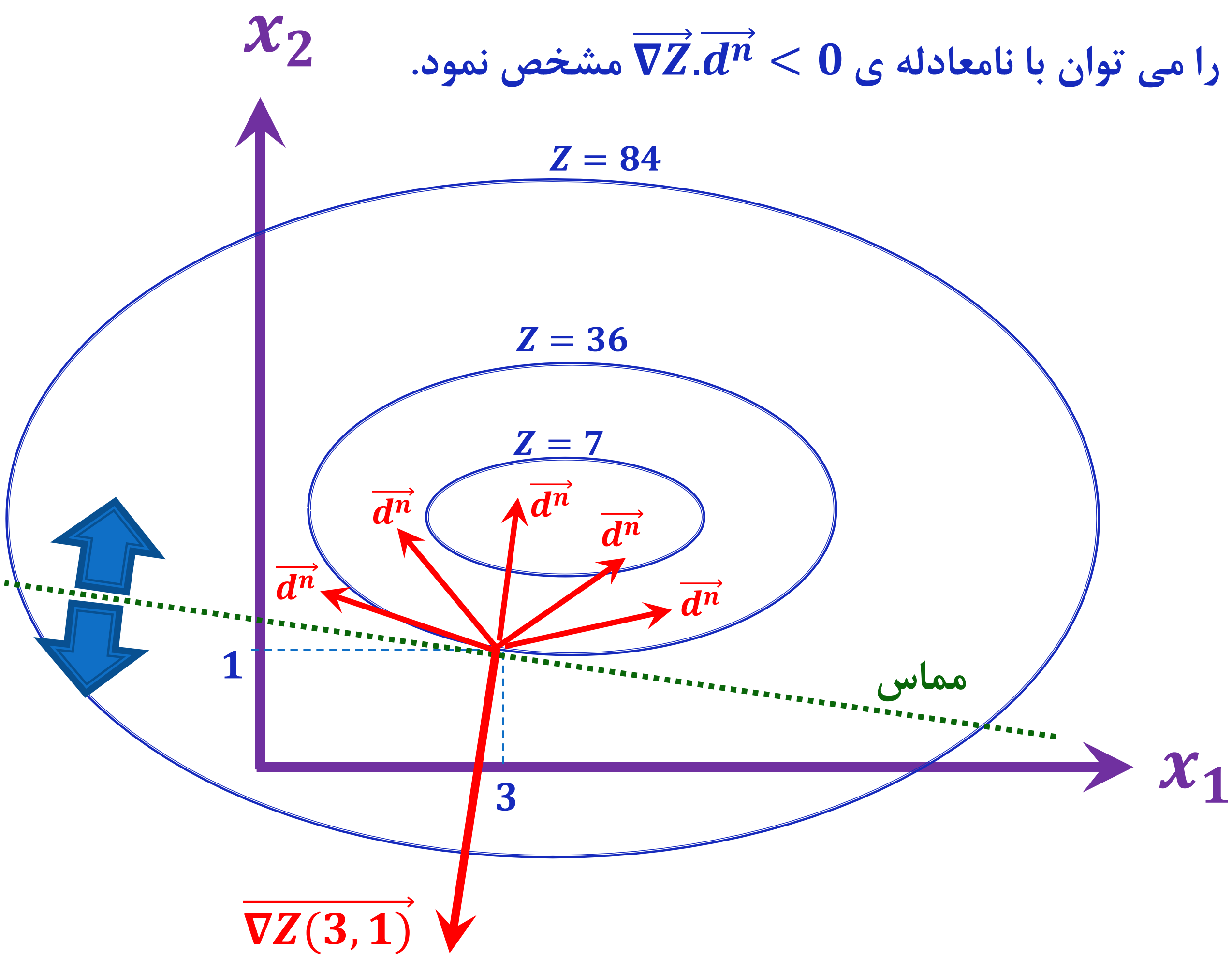
اگر تابع Z (که $Z: R^n \rightarrow R$) در مجموعه نقاط S مشتق پذیر باشد، آن گاه در هر نقطه ی x از S که در آن $\overrightarrow{\nabla Z(x)} \neq \vec{0}$، بردار $\overrightarrow{\nabla Z(x)}$ جهتی را نشان می دهد که در آن جهت، بیشترین افزایش تابع Z وجود دارد.



مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$

- ✓ بردار گرادیان عمود بر خط هم تراز در نقطه x^n است.
- ✓ هریک از جهات کاهشی (فرود $\overrightarrow{d^n}$) را می توان با نامعادله ی $\overrightarrow{\nabla Z}.\overrightarrow{d^n} < 0$ مشخص نمود.



$\overrightarrow{\nabla Z}.\overrightarrow{d^n} < 0$

⇕

کسینوس زاویه بین
دو بردار $\overrightarrow{\nabla Z}$ و $\overrightarrow{d^n}$ منفی باشد.

مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$

بنحوی که: $Z(X^{n+1}) < Z(X^n)$ $\qquad X^{n+1} = X^n + \alpha_n \overrightarrow{d^n}$

سوال: α_n و $\overrightarrow{d^n}$ چه باشند؟

فرض $\begin{cases} \alpha_n = 3\sqrt{2} \\ \overrightarrow{d^n} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \end{cases}$

$$X^{n+1} = (3,1) + 3\sqrt{2}\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = (6,4)$$

$$Z(X^{n+1}) = 0$$

با خوش شانسی، در یک گام به نقطه بهینه رسیدیم!

مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$

بنحوی که: $Z(X^{n+1}) < Z(X^n)$ $\qquad X^{n+1} = X^n + \alpha_n \overrightarrow{d^n}$

سوال: α_n و $\overrightarrow{d^n}$ چه باشند؟

فرض $\overrightarrow{d^n} = (0,1) \Rightarrow X^{n+1} = (3,1) + \alpha_n(0,1) = (3, 1 + \alpha_n)$

باید $Z(x^{n+1}) = Z(3, 1 + \alpha_n) < Z(x^n) = Z(3,1) \rightarrow 9 + 3(\alpha_n - 3)^2 < 9 + 3(9) = 36$

$\rightarrow 0 < \alpha_n < 6$

اگر $\alpha_n > 6$ به نقاط بدتر (بزرگتر) می رسیم.

مثال: یافتن نقطه کمینه تابع با شروع از $X^0 = (3,1)$

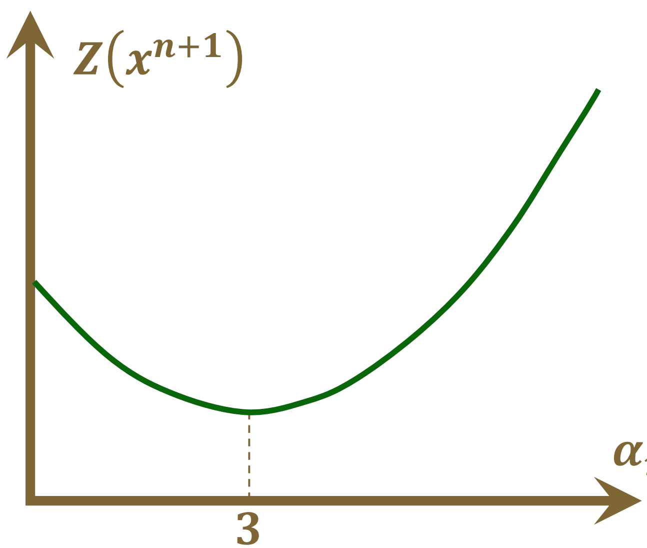
$$Z(x_1, x_2) = (x_1 - 6)^2 + 3(x_2 - 4)^2$$

بنحوی که: $Z(X^{n+1}) < Z(X^n)$ $\qquad$ $X^{n+1} = X^n + \alpha_n \overrightarrow{d^n}$

سوال: α_n و $\overrightarrow{d^n}$ چه باشند؟

$$Z(x^{n+1}) = Z(3, 1 + \alpha_n) = 9 + 3(\alpha_n - 3)^2 < 36$$

یک انتخاب خوب برای α_n : مقداری که عبارت $Z(x^{n+1})$ را کمینه کند.



استفاده از روشهای حل مسائل تک متغیره
(یا در این مثال، مشتق گیری): $\alpha_n = 3$

$\boxed{x^n = (3,1)} \Rightarrow \boxed{x^{n+1} = (3,4)}$

$$Z(X^{n+1}) = 9$$

تکرار روند ...

ضابطه پایانی:

تغییر مقدار تابع کاهش یابد.

$$Z(X^{n-1}) - Z(X^{n}) \leq \kappa$$

مشتق تقریباً به صفر برسد.

$$\underset{i}{Max}\left\{\left|\frac{\partial Z(X^{n})}{\partial x_i}\right|\right\} \leq \kappa$$

درصد تغییر در تکرارها کاهش یابد.

$$\underset{i}{Max}\left\{\frac{\left|x_i^{n} - x_i^{n-1}\right|}{x_i^{n-1}}\right\} \leq \kappa$$

فاصله دو نقطه کاهش یابد

$$\sum_{i}(x_i^{n} - x_i^{n-1})^2 \leq \kappa$$

کمینه سازی بدون محدودیت

تعیین $\overrightarrow{d_n}$

روش تندترین فرود (Steepest Descent):

بصورت حریصانه (greedy) فرود با تندترین شیب انتخاب می شود (تابع هدف در هر تکرار کاهش می یابد).

بنحوی که: $Z(X^{n+1}) < Z(X^n)$ $\qquad$ $X^{n+1} = X^n + \alpha_n \overrightarrow{d^n}$

$$\overrightarrow{d^n} = -\overrightarrow{\nabla Z(X^n)}$$

$$X^{n+1} = X^n - \alpha_n \nabla Z(X^n)$$

$$\nabla Z(X^n) = \left[\frac{\partial Z(X^n)}{\partial x_1}, \frac{\partial Z(X^n)}{\partial x_2}, \dots, \frac{\partial Z(X^n)}{\partial x_I}\right]$$

اگر $\frac{\partial Z(X^n)}{\partial x_i}$ قابل محاسبه مستقیم نباشد:

$$\frac{\partial Z(X^n)}{\partial x_i} \cong \frac{Z(\dots, x_{i-1}^n, x_i^n + \Delta x_i, x_{i+1}^n, \dots) - Z(\dots, x_i^n, \dots)}{\Delta x_i}$$

کمینه سازی بدون محدودیت

تعیین α_n (بزرگی گام)

در جهت تندترین فرود ($-\overrightarrow{\nabla Z(X^n)}$) چه قدر حرکت کنیم؟

$$\begin{aligned}&Min\ Z(X^{n+1}) = Z(X^n - \alpha_n \nabla Z(X^n))\\&s.t.\\&\alpha_n > 0\end{aligned}$$

$$\frac{d}{d\alpha} Z\left(X^n - \alpha_n \nabla Z(X^n)\right) = 0$$

کمینه سازی بدون محدودیت

روش تندترین فرود به کمینه محلی همگرا می شود.

تابع اکیداً محدب باشد: روش تند ترین فرود به کمینه جهانی همگرا می شود.

کمینه سازی بدون محدودیت

قضیه: در روش تندترین فرود جهات حرکت در دو تکرار متوالی بر هم عمود هستند.

$$\overrightarrow{d^n}.\overrightarrow{d^{n+1}} = 0$$

اثبات: برای یافتن α ، قرار دادیم:

$$\frac{d}{d\alpha}Z(X^{n+1}) = 0$$

$$\frac{dZ(X^{n+1})}{d\alpha} = \frac{dZ(X^{n+1})}{dX} \times \frac{dX}{d\alpha} =$$

$$= \left[\frac{\partial Z(X^{n+1})}{\partial x_1}, \frac{\partial Z(X^{n+1})}{\partial x_2}, \dots, \frac{\partial Z(X^{n+1})}{\partial x_1}\right] \times \begin{bmatrix} \frac{dx_1^{n+1}}{d\alpha} \\ \cdot \\ \cdot \\ \cdot \\ \frac{dx_I^{n+1}}{d\alpha} \end{bmatrix} =$$

$$= \left[\frac{\partial Z(X^{n+1})}{\partial x_1}, \frac{\partial Z(X^{n+1})}{\partial x_2}, \dots, \frac{\partial Z(X^{n+1})}{\partial x_1}\right] \times \begin{bmatrix} d_1 \\ \cdot \\ \cdot \\ \cdot \\ d_I \end{bmatrix} = -d^{n+1}.d^n = 0$$

کمینه سازی بدون محدودیت

قضیه: در روش تندترین فرود جهات حرکت در دو تکرار متوالی بر هم عمود هستند.

$$\overrightarrow{d^n}.\overrightarrow{d^{n+1}} = 0$$

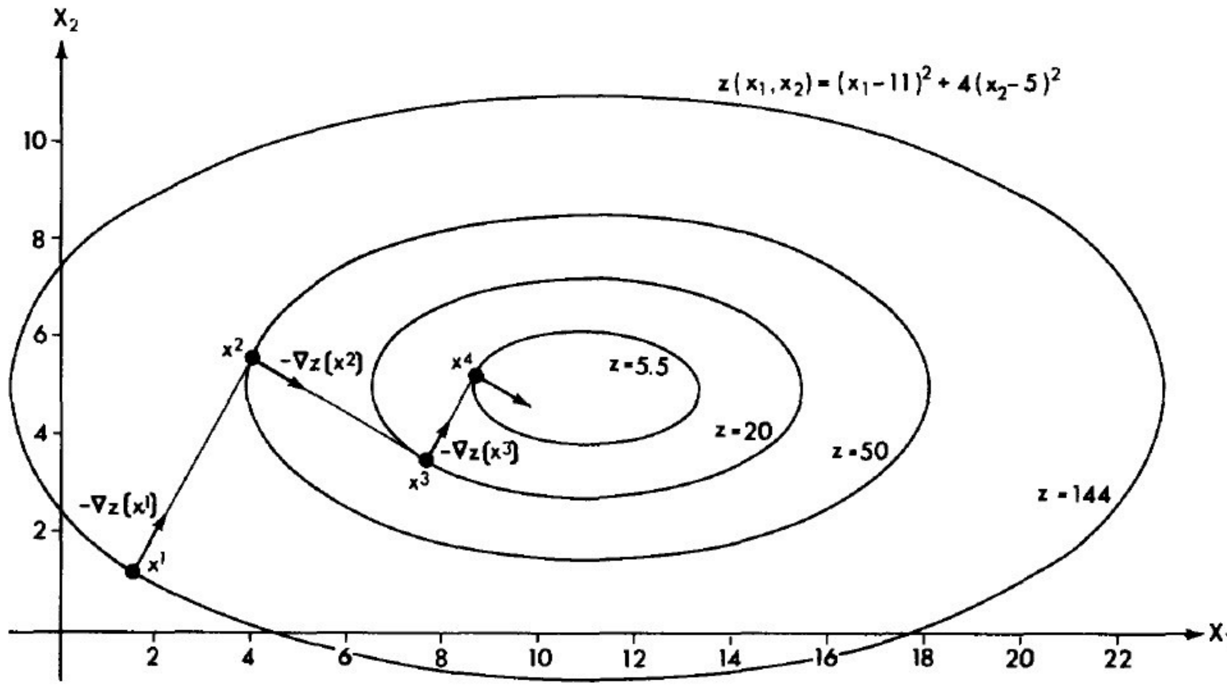


Figure 4.5 Convergence pattern of the steepest descent algorithm; note the zig-zagging of the consecutive descent directions.

کمینه سازی بدون محدودیت

روش نیوتن

استفاده از اطلاعات مشتق دوم (ماتریس هشین)، تقریب Z با استفاده از تقریب درجه ۲ تیلور.

$$X^{n+1} = X^{n} - \nabla Z(X^{n}).\,[H(X^{n})]^{-1}$$

$$X^{n+1} = X^{n} + \alpha_n \overrightarrow{d^n}$$

روش نیوتن $\overrightarrow{d^n} = -\nabla Z(X^{n}).\,[H(X^{n})]^{-1}$ $\quad \alpha_n = 1$

روش تندترین فرود $\overrightarrow{d^n} = -\nabla Z(X^{n})$ $\quad$ محاسبه α_n

روش نیوتن کاربرد زیادی ندارد (نیاز به ماتریس هشین می باشد).

Sheffi Y (1985), Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods, Prentice Hall, New Jersey. (Chapter 4)