Optimizasyon 3

Hesaplamalı Matematik I - Özgür Martin

Kısıtsız Çok Değişkenli Optimizasyon

Dışbükey $f\colon \mathbb{R}^n \to \mathbb{R}$ fonksiyonunun birinci dereceden kısmi türevleri var ve sürekli olsun

Gradyan vektörü:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

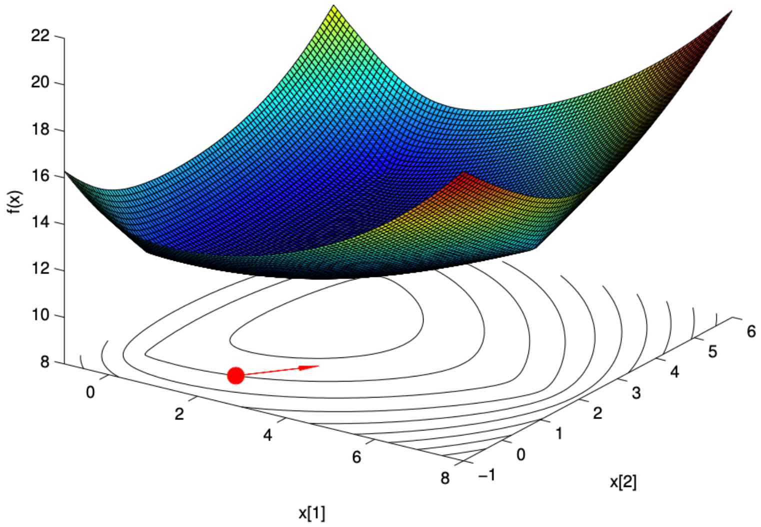
Hatırlatma: f fonksiyonun bir $x \in \mathbb{R}^n$ noktasında en hızlı azaldığı yön gradyanın ter yönüdür, yani $-\nabla f(x)$.

Gradyan İnişi Yöntemi

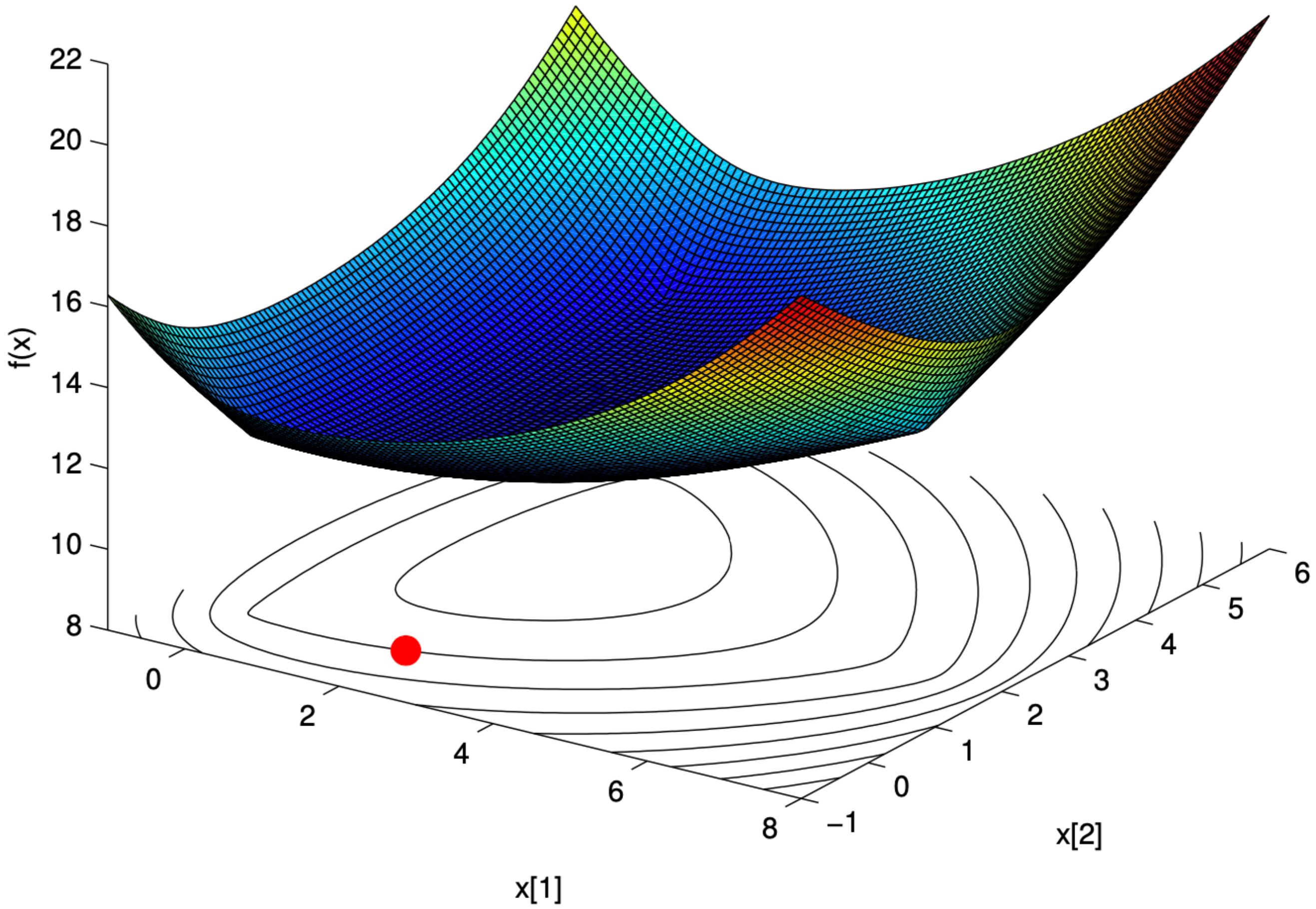
Dışbükey, bütün kısmi türevleri var ve sürekli olan bir $f\colon \mathbb{R}^n \to \mathbb{R}$ fonksiyonu için $\nabla f(x) = 0$ eşitliğini sağlayan $x \in \mathbb{R}^n$ noktasını arıyoruz.

Gradyan İnişi Yöntemi

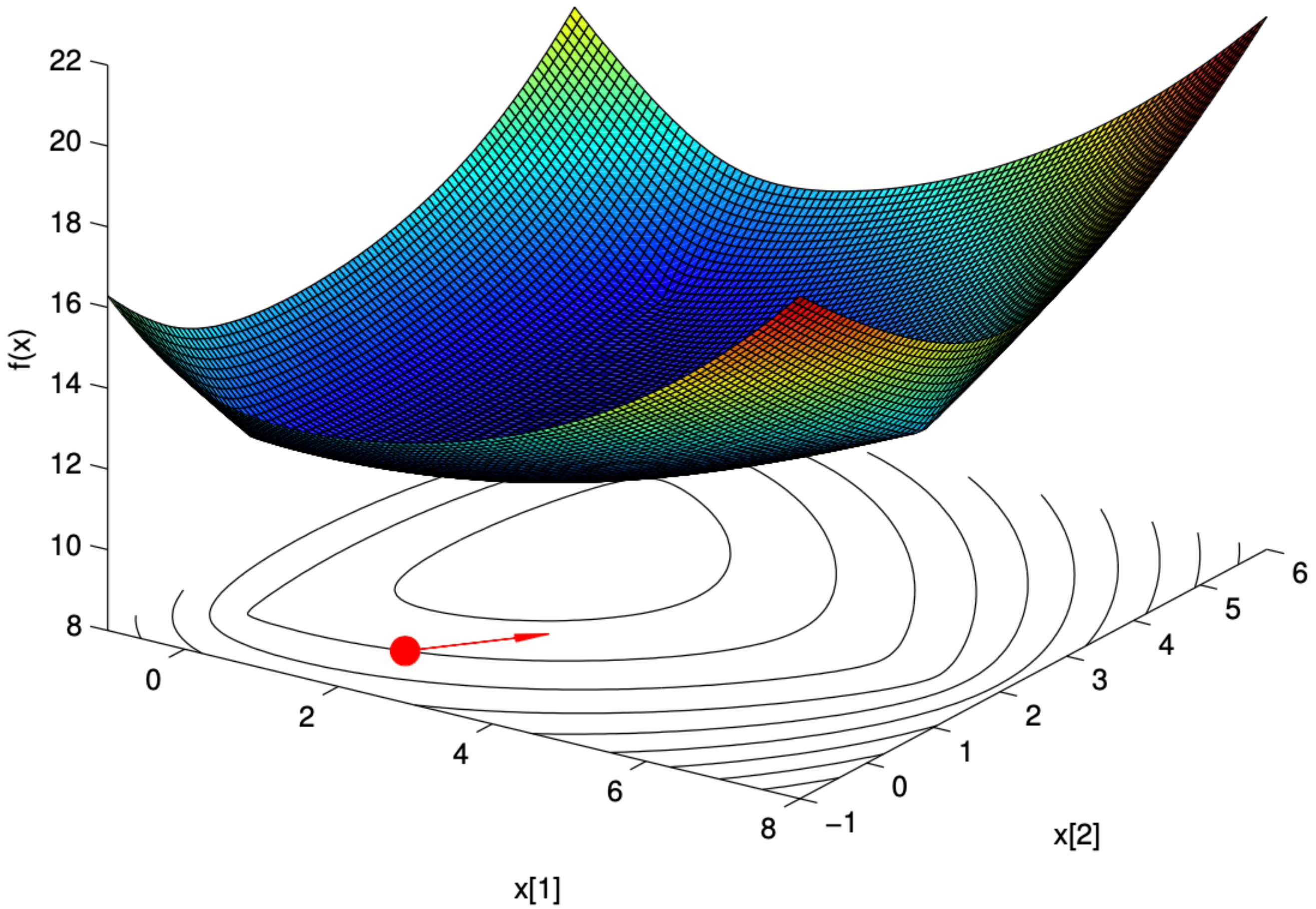
1. Herhangi bir $x \in \mathbb{R}^n$ noktasından başla.
2. Bir ögrenme hızı c seç.
3. $x = x - c\,\nabla f(x)$
4. $\| \nabla f(x) \|$ yeterince küçük olana kadar devam et.



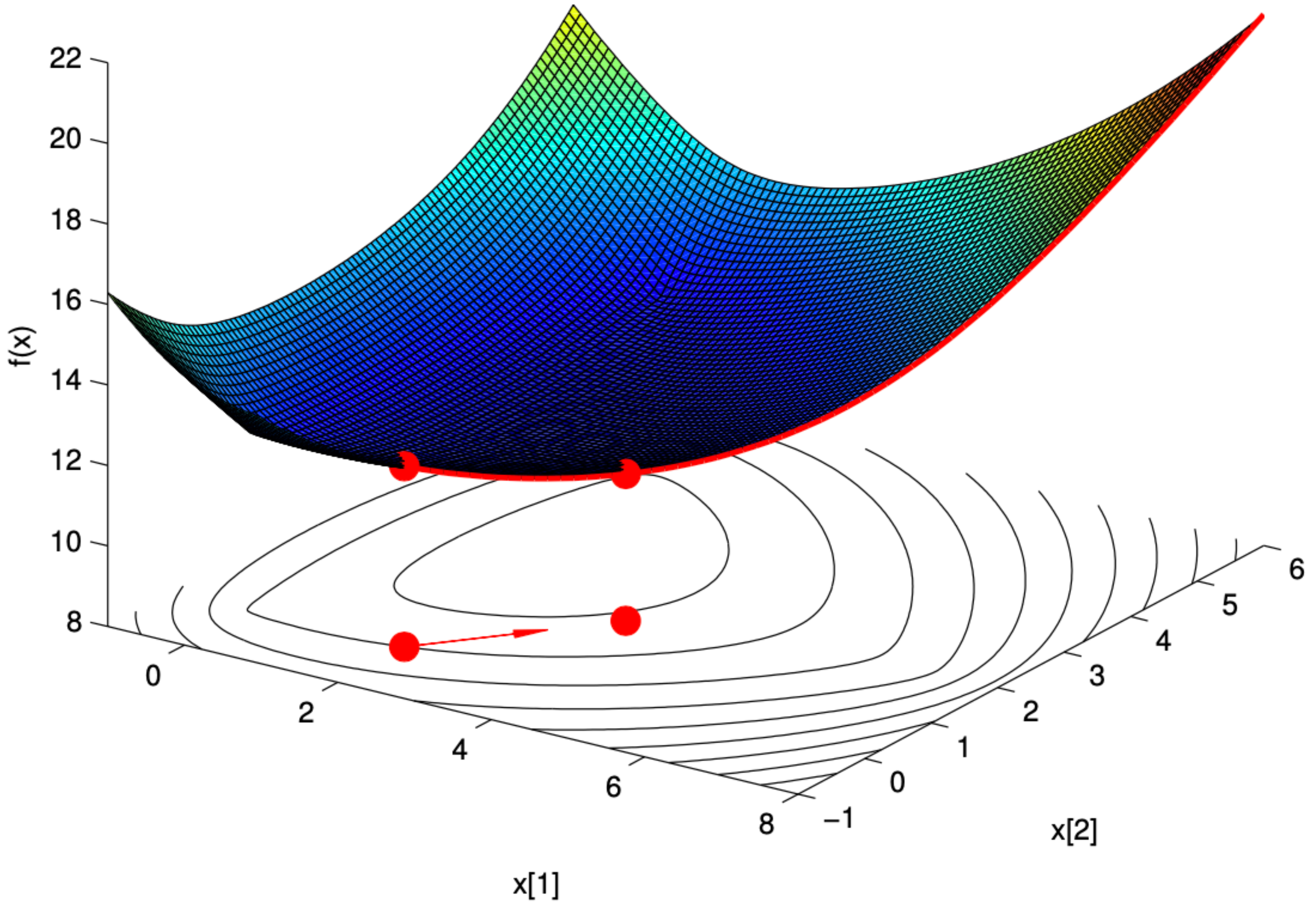
Gradyan İnişi Yöntemi (Çok Değişken)



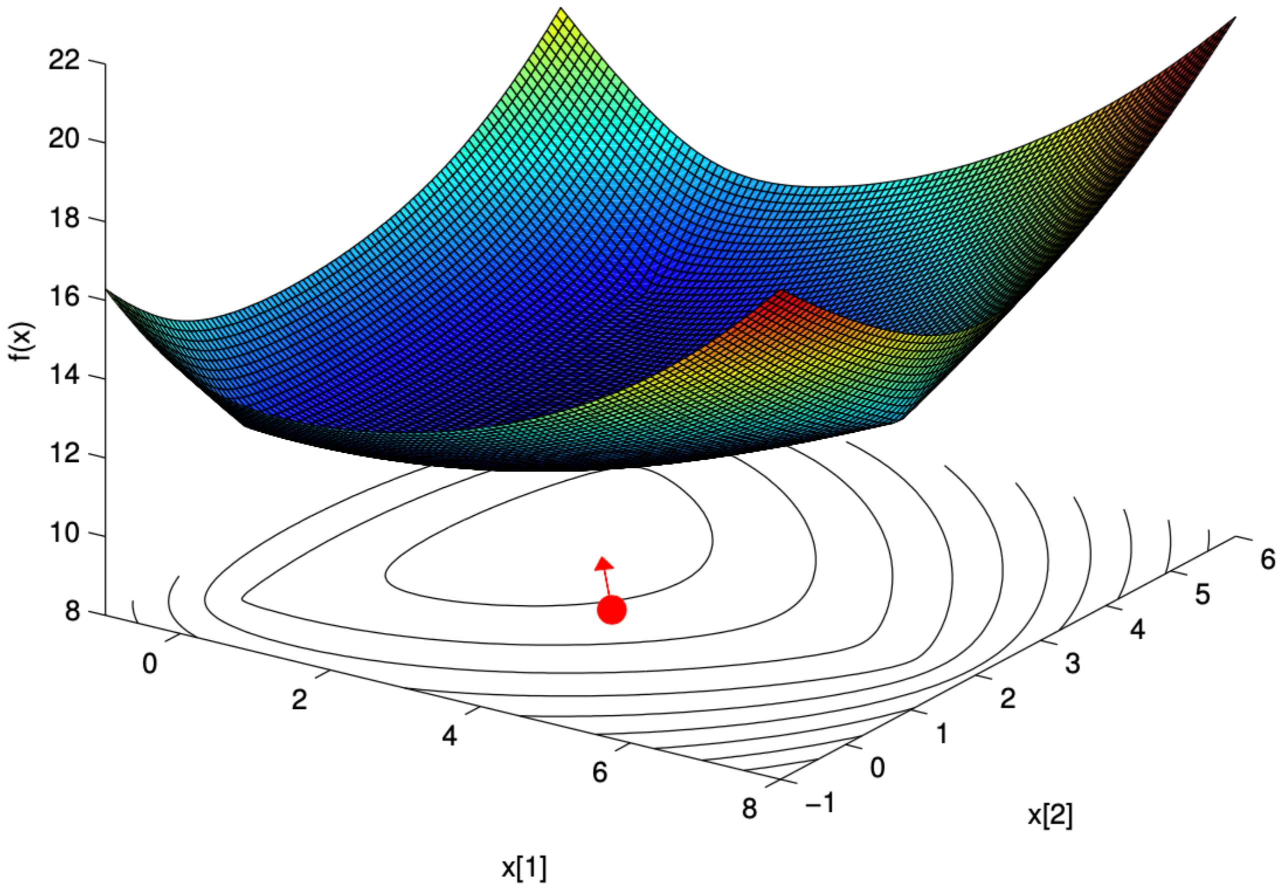
Gradyan İnişi Yöntemi (Çok Değişken)



Gradyan İnişi Yöntemi (Çok Değişken)



Gradyan İnişi Yöntemi (Çok Değişken)

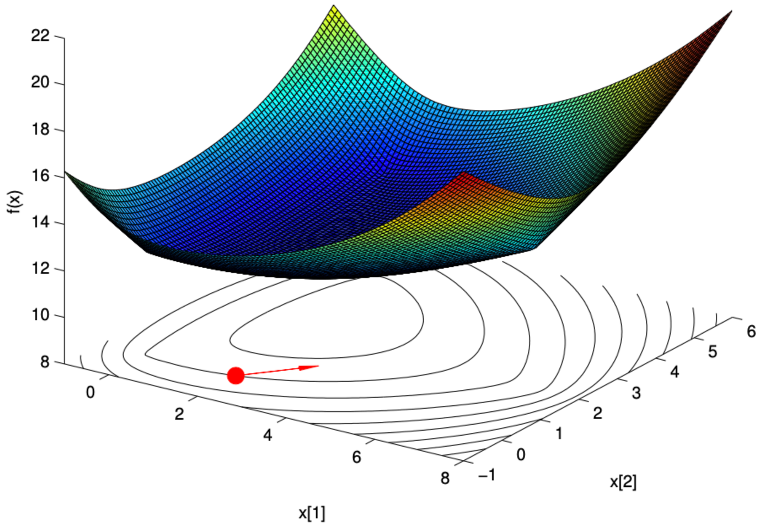


Gradyan İnişi Yöntemi

Dışbükey, bütün kısmi türevleri var ve sürekli olan bir $f: \mathbb{R}^n \to \mathbb{R}$ fonksiyonu için $\nabla f(x) = 0$ eşitliğini sağlayan $x \in \mathbb{R}^n$ noktasını arıyoruz.

Gradyan İnişi Yöntemi

$$x^{(k)} = x^{(k-1)} - c\,\nabla f(x^{(k-1)})$$



Gradyan İnişi Yöntemi

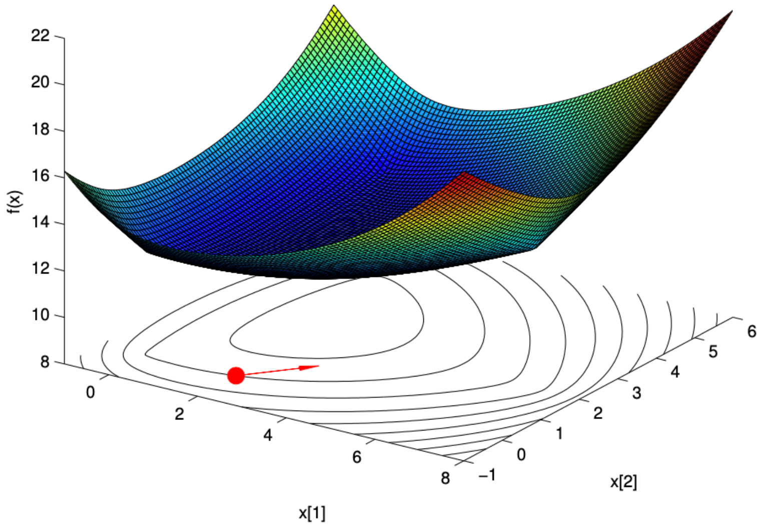
Dışbükey, bütün kısmi türevleri var ve sürekli olan bir $f\colon \mathbb{R}^n \to \mathbb{R}$ fonksiyonu için $\nabla f(x) = 0$ eşitliğini sağlayan $x \in \mathbb{R}^n$ noktasını arıyoruz.

Gradyan İnişi Yöntemi

İstersek öğrenme hızını her adımda değiştirebiliriz: ***c*** yerine $\boldsymbol{c_k}$

$$x^{(k)} = x^{(k-1)} - c_k \nabla f(x^{(k-1)})$$

Mesela: $c_k = \frac{1}{k}$ veya $c_k = \frac{1}{\sqrt{k}}$



Rassal Gradyan İnişi Yöntemi

Problem:

$$\min_{x} \frac{1}{m}\sum_{i=1}^{m} f_i(x)$$

Uygulamada *m* veri noktası sayısı. Mesela KKT.

Burada $\nabla \sum_{i=1}^{m} f_i(x) = \sum_{i=1}^{m} \nabla f_i(x)$

Gradyan İniş Yöntemi:

$$x^{(k)} = x^{(k-1)} - t_k \cdot \frac{1}{m}\sum_{i=1}^{m} \nabla f_i(x^{(k-1)}), \quad k = 1, 2, 3, \ldots$$

Her adımda *m* tane gradyan hesabı yapmamız gerekiyor. Büyük veriler için çok masraflı!

Rassal Gradyan İnişi Yöntemi

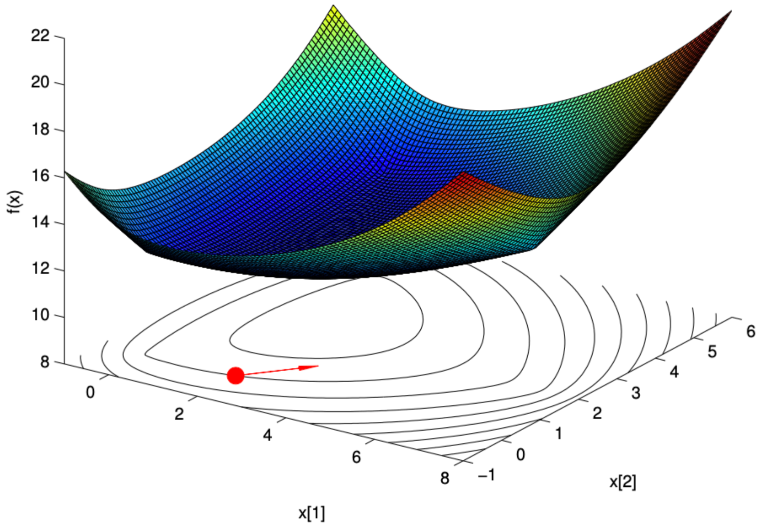
Problem:

$$\min_{x} \frac{1}{m} \sum_{i=1}^{m} f_i(x)$$

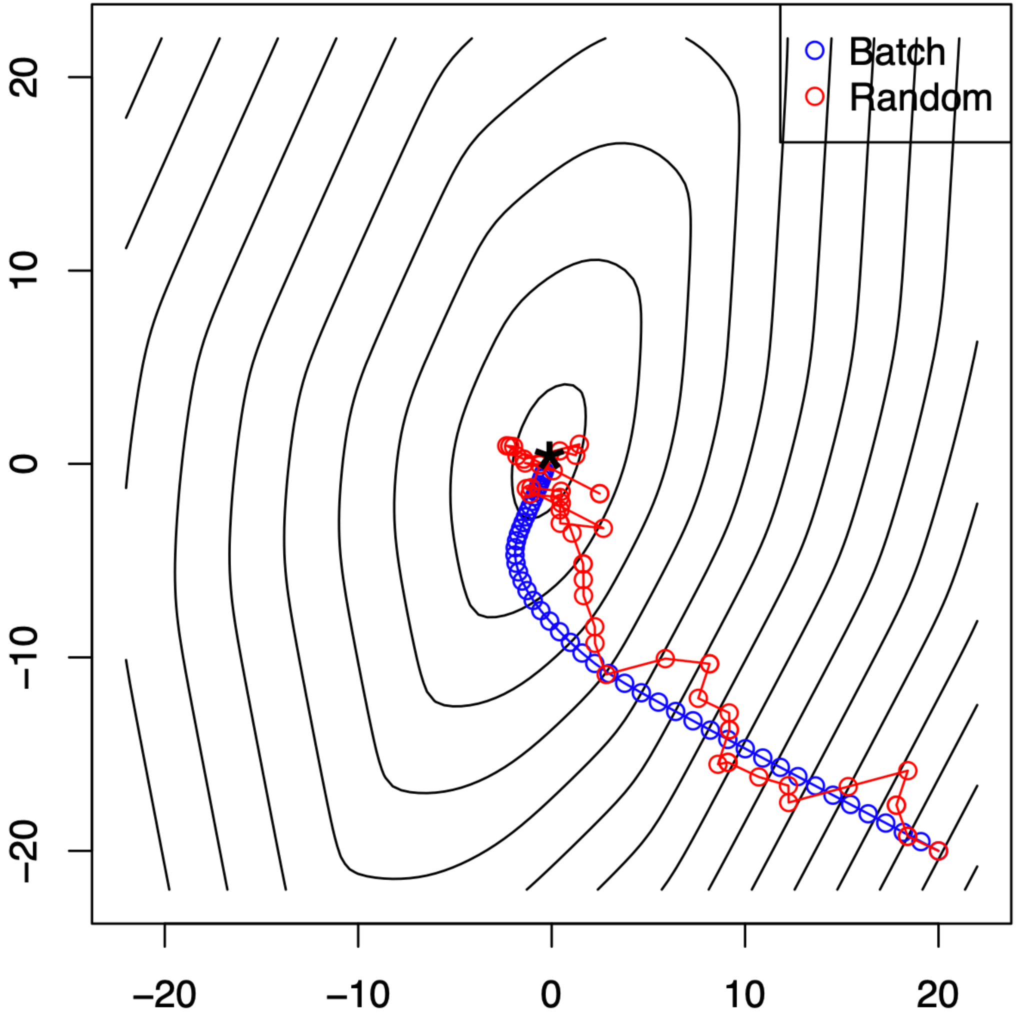
Rassal Gradyan İnişi Yöntemi

$$x^{(k)} = x^{(k-1)} - c_k \cdot \nabla f_{i_k}(x^{(k-1)})$$

burada $i_k \in \{1, \ldots, m\}$ her adımda düzgün dağılıma göre rasgele seçilsin.



Rassal Gradyan İnişi Yöntemi



Mavi: Gradyan İniş
Kırmızı: Rassal Gradyan İniş

RGİ genellikle

- minimum değerden uzakken iyi çalışır
- minimuma yakınken zorlanır

Rassal Gradyan İnişi Yöntemi

Yakınsaklık analizleri (c_k) adım uzunluğu dizisinin şu şartları sağlaması gerektiğini göstermiştir:

$$c_k \xrightarrow[k\to\infty]{} 0 \qquad \text{ve} \qquad \sum_{k=1}^{\infty} c_k = \infty.$$

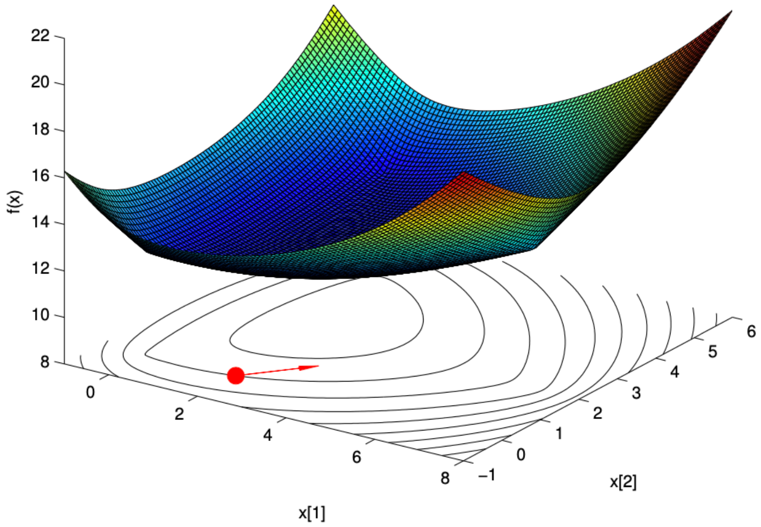
Uygulamada $\quad c_k = \dfrac{\epsilon}{\sqrt{k}} \quad$ ve $\quad \epsilon = 10^{-4}$.

Rassal Gradyan İnişi Yöntemi (Genelleşmiş)

Problem: $$\min_x \frac{1}{m}\sum_{i=1}^{m} f_i(x)$$

Rassal Gradyan İnişi Yöntemi

$$x^{(k)} = x^{(k-1)} - c_k \cdot \sum_{i\in K} \nabla f_i(x^{(k-1)})$$



burada $\mathrm{K} \subset \{1, \dots, m\}$ rasgele seçilmiş.

Rassal Gradyan İnişi Yöntemi

Ne zaman duracağız?

- Fonksiyonun değerindeki düşüşler küçülünce
- $x^{(k)}$ noktasındaki değişimler azalınca