تخمین و شناسایی و سیستم ها

Estimation & System Identification

بابک نجار اعرابی

دانشکده مهندسی برق و کامپیوتر دانشگاه تهران

نیم سال دوم سال تحصیلی 1400-01

موضوع این جلسه

مروری بر روش های بهینه سازی

جلسه دوم

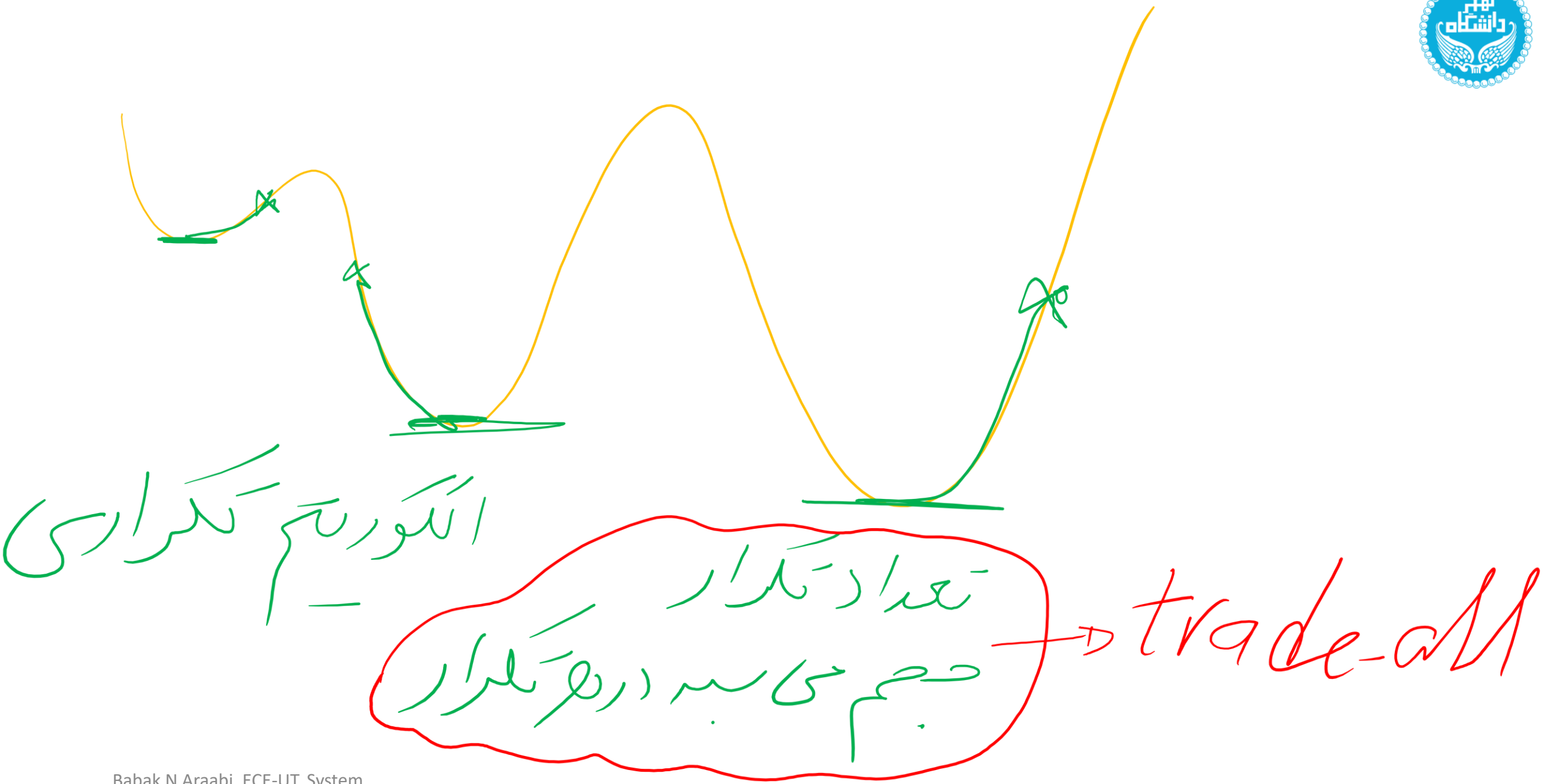
Babak Nadjar Araabi
School of Electrical & Computer Eng
University of Tehran

ECE-UT - Spring 2022

Optimization

- Local opt.
- Global opt. → Exhaustive search
  → Random search

local optimum


الگوریتم تکراری
تعداد تکرار
حجم می‌سبد در هر تکرار
trade-all

Direct Search Methods

$I(\underline{\theta})$ $\underline{\theta}^* = \arg\min_{\underline{\theta}} I(\underline{\theta})$

روش‌هایی که از $I(\underline{\theta})$ برای یافتن $\underline{\theta}$ استفاده می‌کنند

- Simplex search (Nelder-Mead Method)
- Hooke & Jeeves Method

Gradient-based search Methods

⊛ steepest descent

بسط تیلور حول $\underline{\theta}_0$

$$I(\underline{\theta}) = I(\underline{\theta}_0) + (\underline{\theta} - \underline{\theta}_0)^T \nabla I\Big|_{\underline{\theta}=\underline{\theta}_0} + \frac{1}{2}(\underline{\theta} - \underline{\theta}_0)^T \nabla^2 I\Big|_{\underline{\theta}=\underline{\theta}_0} (\underline{\theta} - \underline{\theta}_0) + \cdots$$

$$\nabla I = \frac{\partial I}{\partial \underline{\theta}} = \left[\frac{\partial I}{\partial \theta_i}\right]_{n\times 1} \qquad \nabla^2 I = \frac{\partial^2 I}{\partial \underline{\theta}^2} = \left[\frac{\partial^2 I}{\partial \theta_i \partial \theta_j}\right]_{n\times n}$$

تقریب خطی تابع $J(\theta)$ حول θ_0

$$J(\theta) = J(\theta_0) + (\theta - \theta_0)^T \nabla J\Big|_{\theta=\theta_0}$$

$J(\theta) < J(\theta_0)$ منفی

$\theta - \theta_0 \propto \nabla J(\theta_0)$

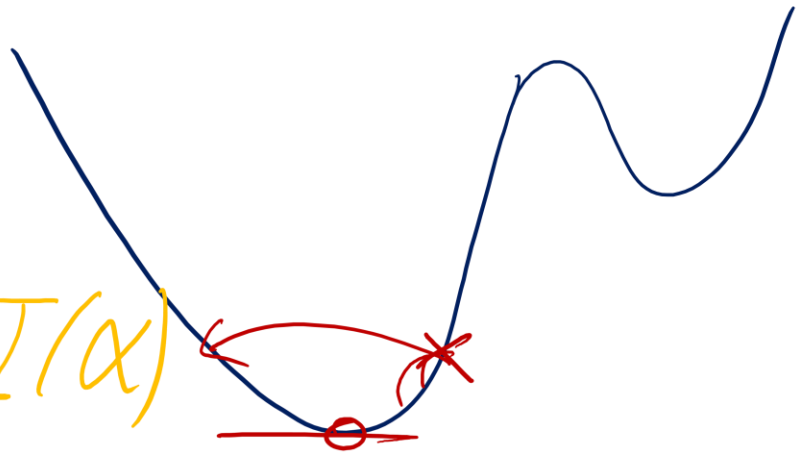
$$\theta_1 - \theta_0 = -\alpha \nabla J(\theta_0)$$

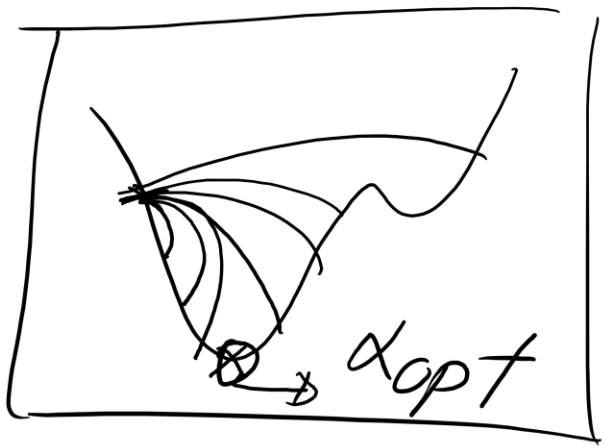
$$\theta_1 = \theta_0 - \alpha \nabla J(\theta_0)$$

$$I(\underline{\theta}_1) = I(\underline{\theta}_0) - \alpha \|\nabla I(\underline{\theta}_0)\|^2$$

step size α

$$I(\theta_1) = I\left(\underbrace{\underline{\theta}_0 - \alpha \nabla I(\underline{\theta}_0)}_{\underline{\theta}_1}\right) = I(\alpha)$$

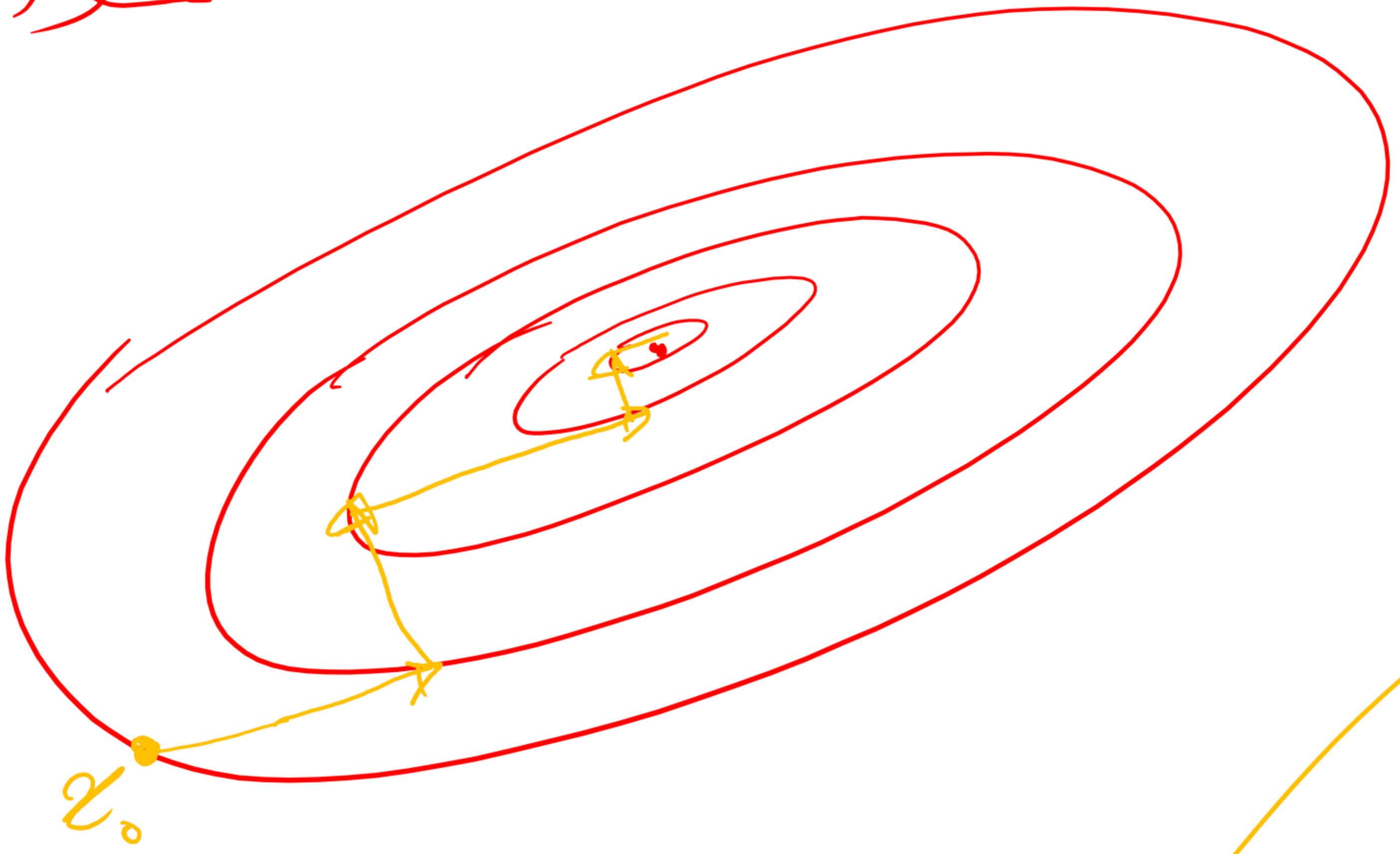


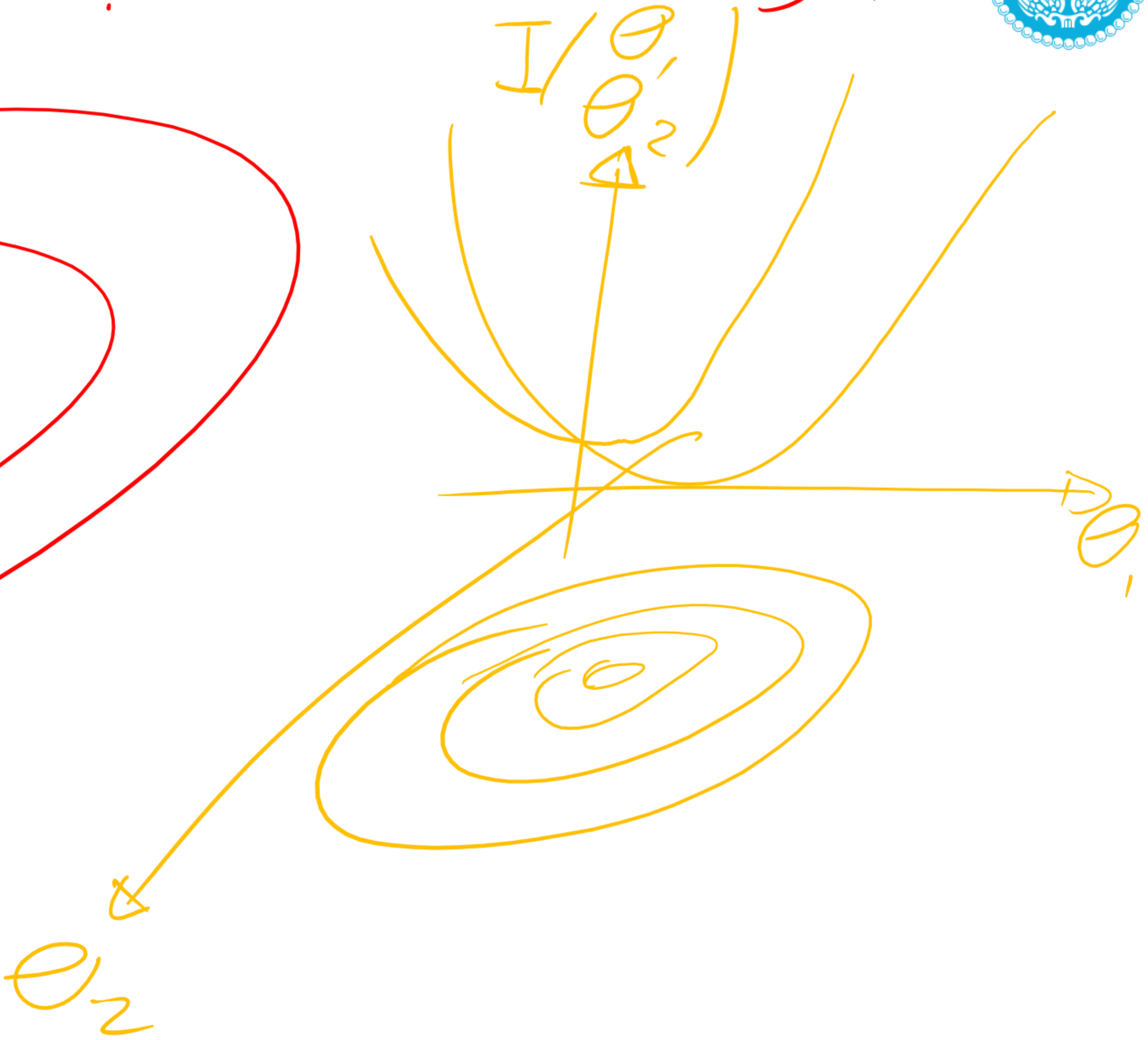


α تعیین step size به روش Line search

$$\frac{\partial I}{\partial \alpha} = \underbrace{\nabla I(\underline{\theta}_1)^T}_{\text{وقت گام ۲}} \underbrace{\nabla I(\underline{\theta}_0)}_{\text{وقت گام ۱}} = 0$$

اگر Line Search انجام شود به جهات حرکت بریم
کسوانه





Newton Method

$$J(\underline{\theta}) = J(\underline{\theta}_0) + (\underline{\theta} - \underline{\theta}_0)^T \nabla J(\underline{\theta}_0) + \frac{1}{2}(\underline{\theta} - \underline{\theta}_0)^T \nabla^2 J(\underline{\theta}_0)(\underline{\theta} - \underline{\theta}_0) + \dots$$

اگر $\underline{\theta}_1$ کمینه $J(\underline{\theta})$ باشد، آنگاه $\underline{\underline{H}}_J(\underline{\theta}_0)$

$$\nabla J(\underline{\theta}_1) = \underline{0} \rightarrow \nabla J(\underline{\theta}_0) + \underbrace{\nabla^2 J(\underline{\theta}_0)}_{\underline{\underline{H}}_J(\theta_0)} (\underline{\theta}_1 - \underline{\theta}_0) = \underline{0}$$

$$\underline{\theta}_1 = \underline{\theta}_0 - \underline{\underline{H}}_J^{-1}(\underline{\theta}_0) \nabla J(\underline{\theta}_0)$$

Quasi-Newton Method

$$\underline{\theta}_{i+1} = \underline{\theta}_i - \alpha_i \underline{\underline{S}}_i \underline{g}_i$$

step size ← α_i , $\underline{\underline{S}}_i$, $\underline{g}_i$ → $\nabla I(\underline{\theta}_i)$

$\underline{\underline{S}}_i$ تقریبی از معکوس ماتریس بی وارون هسین $\underline{\underline{H}}_I^{-1}(\underline{\theta}_i)$

BFGS

$$\underline{\underline{S}}_{i+1} = \underline{\underline{S}}_i + \underline{\underline{Q}}_i \qquad \underline{\underline{S}}_0 = \underline{\underline{I}}$$

$$\underline{\underline{S}}_{i+1} = \left(\underline{\underline{I}} - \frac{\Delta\underline{\theta}_i \cdot \Delta\underline{g}_i^T}{\Delta\underline{\theta}_i^T \Delta\underline{g}_i}\right) \underline{\underline{S}}_i \left(\text{متقارن بودن}\right)^T + \frac{\Delta\underline{x}_i \cdot \Delta\underline{x}_i^T}{\Delta\underline{x}_i^T \Delta\underline{g}_i}$$

BFGS

$$\Sigma_i > 0 \Rightarrow \Sigma_{i+1} > 0$$

اگر $\Sigma_0 = I$ و تابع quadratic
همگرایی در n گام تضمین شده است

$$\Sigma_n = H_I^{-1}(\theta^*)$$