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 سوم

Babak Nadjar Araabi

School of Electrical & Computer Eng

University of Tehran

ECE-UT - Spring 2022

Quasi-Newton Method

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

stepsize

$\underline{I}(\underline{\theta})$ کمینه‌جو

$$\underline{\underline{H}}^{-1}(\underline{\theta}_i) = \underline{\underline{\nabla}}^2 I(\underline{\theta})\Big|_{\underline{\theta}=\underline{\theta}_i}$$

$$\underline{\underline{I}}$$

$$\nabla I(\underline{\theta})\Big|_{\underline{\theta}=\underline{\theta}_i}$$

$\underline{\underline{S}}_0 = \underline{\underline{I}}$ $\quad$ $\underline{\underline{S}}_i > 0 \Rightarrow \underline{\underline{S}}_{i+1} > 0$

$O(n^2)$

Conjugate Gradient Method

$$\underline{\theta}_{i+1} = \underline{\theta}_i - \alpha_i \underline{P}_i \qquad O(n)$$

$$\underline{P}_i = \underline{g}_i - \beta_i \underline{P}_{i-1}$$

$$\underline{g}_i = \nabla J(\underline{\theta})\Big|_{\underline{\theta}=\underline{\theta}_i} \qquad P_0 = g_0$$

$$\beta_i = \frac{\underline{g}_i^T \underline{g}_i}{\underline{g}_{i-1}^T \underline{g}_{i-1}} = \frac{\|\underline{g}_i\|^2}{\|\underline{g}_{i-1}\|^2}$$

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

Fletcher, R. , 198a

Practical Methods of Optimization

Vol I: Unconstrained Optimization

$$I(\underline{\theta}) = \sum_{i=1}^{N} e^2(i) \qquad e(i) = y(i) - \hat{y}(i)$$

$$e(i;\underline{\theta}) = y(i) - \underline{x}^T(i)\underline{\theta}$$

$$\to f(i;\underline{\theta})$$

Least Squares

$$I(\underline{\theta}) = \sum_{i=1}^{N} f^2(i;\underline{\theta})$$

Nonlinear LS

گرادیان = ?　　هسین ?

$$J(\underline{\theta}) = \sum_{i=1}^{N} f^2(i;\underline{\theta}) = \underline{f}^T \underline{f}$$

$$\frac{\partial J}{\partial \underline{\theta}} \qquad \underline{\underline{\nabla}}^2 J(\underline{\theta})$$

$$\frac{\partial J}{\partial \underline{\theta}} = \nabla J(\underline{\theta}) = \underline{g} = \left(\frac{\partial J}{\partial \theta_1}, \frac{\partial J}{\partial \theta_2}, \dots, \frac{\partial J}{\partial \theta_n}\right)^T$$

$$\frac{\partial J}{\partial \theta_j} = 2 \sum_{i=1}^{N} f(i;\underline{\theta}) \frac{\partial f(i;\underline{\theta})}{\partial \theta_j} \qquad j = 1, 2, \dots, n$$

$$I(\underline{\theta}) = \underline{f}^T \underline{f} = \sum_{i=1}^{N} f^2(i;\underline{\theta})$$

$$\underset{N\times 1}{\underline{f}} = \left(f(1,\underline{\theta}), f(2,\underline{\theta}), \ldots, f(N,\underline{\theta})\right)^T$$

$$\underline{\underline{J}} = \frac{\partial \underline{f}}{\partial \underline{\theta}} \triangleq \left[\frac{\partial f(i)}{\partial \theta_j}\right]_{N\times n}$$

ژاکوبین تابع برداری $\underline{f}$
نسبت به بردار $\underline{\theta}$

$$\frac{\partial I}{\partial \underline{\theta}} = 2\,\underline{\underline{J}}^T \underline{f}$$

$$\underline{\underline{H}}\,I(\underline{\theta}) = \underline{\underline{\nabla}}^2 I(\underline{\theta}) \rightarrow H_{ij} = \frac{\partial^2 I(\underline{\theta})}{\partial \theta_i \, \partial \theta_j}$$

$$\underbrace{\frac{\partial}{\partial \theta_i}\left(2\sum_{\ell=1}^{N} f(\ell)\, \frac{\partial f(\ell,\underline{\theta})}{\partial \theta_j} \right)}_{\frac{\partial}{\partial \theta_i}\left(\frac{\partial I}{\partial \theta_j}\right)} = 2\sum_{\ell=1}^{N} \frac{\partial f(\ell)}{\partial \theta_i}\,\frac{\partial f(\ell)}{\partial \theta_j} + 2\sum_{\ell=1}^{N} f(\ell)\,\frac{\partial^2 f(\ell)}{\partial \theta_i \, \partial \theta_j}$$

$$= 2\,\underline{\underline{J}}^T \underline{\underline{J}} + 2\sum_{\ell=1}^{N} f(\ell)\, \underline{\underline{\nabla}}^2 f(\ell)$$

$$\underline{\underline{H}}_I(\underline{\theta}) = \underline{\underline{\nabla}}^2 I(\underline{\theta}) =$$

$$2\,\underline{\underline{J}}^T \underline{\underline{J}} + 2\sum_{l=1}^{N} f(l)\,\underline{\underline{\nabla}}^2 f(l)$$

صرف نظر

فرض کوچک بودن

$(f(l))$

$$\underline{\underline{H}} \simeq 2\,\underline{\underline{J}}^T \underline{\underline{J}} \rightarrow$$ روش گوس نیوتن

$$\underline{\theta}_{i+1} = \underline{\theta}_i - \alpha_i \underbrace{\underline{H}_I^{-1}(\underline{\theta}_i)}_{\approx (2\underline{\underline{J}}^T\underline{\underline{J}})^{-1}} \underbrace{\underline{\nabla}_I(\underline{\theta}_i)}_{2\underline{\underline{J}}^T\underline{f}}$$

$$\boxed{\underline{\theta}_{i+1} = \underline{\theta}_i - \alpha_i (\underline{\underline{J}}_i^T \underline{\underline{J}}_i)^{-1} \underline{\underline{J}}_i^T \underline{f}_i}$$

↳ Gauss-Newton method

Levenberg-Marquardt Method

Regularized version of Gauss-Newton Method

$$\underline{\theta}_{i+1} = \underline{\theta}_i - \alpha_i \left(\underline{\underline{J}}_i^T \underline{\underline{J}}_i + \gamma_i \underline{\underline{I}} \right)^{-1} \underline{\underline{J}}_i^T \underline{f}_i$$

M. B. Menhaj دکتر محمدباقر منهاج