



KTH Computer Science
and Communication

Föreläsning 6

Ickelinjär modellanpassning

Vi har tidigare sett hur man kan lösa överbestämda linjära ekvationssystem $A\mathbf{x} \approx \mathbf{b}$ i minstakvadratmening genom att lösa normalekvationerna $A^T A\mathbf{x} = A^T \mathbf{b}$. Vi ska här visa hur man kan hantera motsvarande överbestämda *olinjära* problem,

$$\mathbf{F}(\mathbf{x}) \approx 0, \quad (1)$$

där

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m, \quad \mathbf{F}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \dots, x_m) \\ \vdots \\ f_n(x_1, \dots, x_m) \end{pmatrix} \in \mathbb{R}^n, \quad n > m.$$

Notera att antalet ekvationer n är större än antalet obekanta m . För att lösa (1) i minstakvadratmening kan man använda *Gauss-Newton's metod*. Formen för metoden är mycket lik Newtons metod för olinjära system av ekvationer. I Newton löser man ekvationssystemet $J\boldsymbol{\delta} = -\mathbf{F}$ för att hitta uppdateringen, $\mathbf{x}_{i+1} = \mathbf{x}_i + \boldsymbol{\delta}$. Här är J jakobianen till $\mathbf{F}$. I Gauss-Newton löser man samma system, men i minstakvadratmening, dvs man löser normalekvationerna, $J^T J\boldsymbol{\delta} = -J^T \mathbf{F}$. Metoden blir

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (J(\mathbf{x}_i)^T J(\mathbf{x}_i))^{-1} J(\mathbf{x}_i)^T \mathbf{F}(\mathbf{x}_i), \quad J(\mathbf{x}) = \{j_{k\ell}\}, \quad j_{k\ell} = \frac{\partial f_k(\mathbf{x})}{\partial x_\ell}.$$

För tillräckligt bra startgissningar konvergerar denna iterativa metod normalt mot lösningen till minimeringsproblemet

$$\min_{\mathbf{x}} \|\mathbf{F}(\mathbf{x})\|_2,$$

dvs minstakvadratlösningen till (1). Se uppgifterna 4.22–4.26 i ENM för exempel. Notera att backslash i MATLAB ger en minstakvadratlösning så snart antal rader i matrisen är större än antal kolumner. Därför blir en implementation av Gauss-Newton i MATLAB så gott som identisk med en implementation av Newtons metod för system.

Härledning

Gauss-Newton kan ses som en approximation av Newtons metod applicerat på problemet

$$\nabla r(\mathbf{x}) = 0, \quad r(\mathbf{x}) := \|\mathbf{F}(\mathbf{x})\|^2, \quad (2)$$

vilket ska vara uppfyllt då $\mathbf{x}$ minimerar $\|\mathbf{F}(\mathbf{x})\|$. Newtons metod för (2) blir

$$\mathbf{x}_{i+1} = \mathbf{x}_i - H(\mathbf{x}_i)^{-1} \nabla r(\mathbf{x}_i), \quad H = \{h_{k\ell}\}, \quad h_{k\ell} = \frac{\partial^2 r}{\partial x_k \partial x_\ell},$$

där $H(\mathbf{x})$ är Hessianen för $r(\mathbf{x})$. Vi noterar att $r = \mathbf{F}^T \mathbf{F}$. Efter lite algebra får man att,

$$\nabla r = 2J^T \mathbf{F}, \quad H = 2J^T J + 2 \sum_{j=1}^n f_j H_j, \quad H_j = \{h_{k\ell}\}, \quad h_{k\ell} = \frac{\partial^2 f_j}{\partial x_k \partial x_\ell}.$$

I Gauss–Newton approximerar man alltså den andra termen i uttrycket för H med noll, summerar över $f_j H_j \approx 0$. När detta är en bra approximation, tex då $\mathbf{F} \approx 0$ eller då $\mathbf{F}$ ungefär linjär nära minimumet, konvergerar Gauss-Newton kvadratisk. Annars blir konvergensen som i fixpunktfallet linjär.