SF1624 Algebra och geometri

Elfte föreläsningen

Mats Boij

Institutionen för matematik
KTH

16 november, 2009

Inkonsistent eller överbestämt?

- När det saknas lösning har vi tidigare kallat ekvationssystemet inkonsistent.
- I tillämpningar är det ofta naturligt att det saknas exakta lösningar och vi kallar detta för överbestämda ekvationssystem.

Exempel

Om vi har mätt upp tre mätvärden för ström och spänning över en resistans kan vi ställa upp ekvationssystemet

$$\begin{cases} Ri_1 & = u_1 \\ Ri_2 & = u_2 \\ Ri_3 & = u_3 \end{cases}$$

Det finns då troligen inget värde på R som uppfylller alla tre ekvationer.

Minsta-kvadratmetoden

Vi kan försöka så gott vi kan och hitta värden på variablerna som är så nära att vara en lösning som möjligt.

Vad menas med nära?

Vi kan säga att $\overline{x}$ är nära att vara en lösning till $A\overline{x} = \overline{b}$ om avståndet mellan $A\overline{x}$ och $\overline{b}$ är litet.

Exempel

Om vi har de tre ekvationssystemet

$$\begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix} R = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

vill vi minimera längden av $(Ri_1 - u_1, Ri_2 - u_2, Ri_3 - u_3)$. Vi kan se det som att hitta den punkt på linjen $t(i_1, i_2, i_3)$ som ligger närmast punkten (u_1, u_2, u_3).

fortsatt exempel

Vi får kortast avstånd när
$(Ri_1 - u_1, Ri_2 - u_2, Ri_3 - u_3) \perp (i_1, i_2, i_3)$, dvs när

$$(Ri_1 - u_1)i_1 + (Ri_2 - u_2)i_2 + (Ri_3 - u_3)i_3 = 0$$

vilket är ekvivalent med

$$R(i_1^2 + i_2^2 + i_3^2) = i_1u_1 + i_2u_2 + i_3u_3$$

och

$$R = \frac{i_1u_1 + i_2u_2 + i_3u_3}{i_1^2 + i_2^2 + i_3^2}$$

är det värde på R som ligger närmast en lösning i *minsta-kvadratmening*.

Generaliserat exempel

När vi löste problemet i vårt exempel använde vi tre mätvärden, men de räkningar och resonemang vi gjorde fungerar för n mätvärden. Vi får att

$$R = \frac{i_1 u_1 + i_2 u_2 + \cdots + i_n u_n}{i_1^2 + i_2^2 + \cdots + i_n^2}$$

ligger närmast en lösning till

$$\begin{cases} Ri_1 & = & u_1 \\ Ri_2 & = & u_2 \\ \vdots & \vdots & \vdots \\ Ri_n & = & u_n \end{cases}$$

Obs! Det är inte medelvärdet av $u_1/i_1, u_2/i_2, \ldots, u_n/i_n$.

Två variabler

Om vi har två variabler, x och y, kan vi uttrycka ekvationssystemet som

$$x \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{pmatrix} + y \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \Leftrightarrow x\overline{u} + y\overline{v} = \overline{b}$$

Vi vill på samma sätt komma så nära högerledet som möjligt och då måste skillnaden

$$x\overline{u} + y\overline{v} - \overline{b}$$

vara ortogonal mot planet som spänns upp av $\overline{u}$ och $\overline{v}$, dvs

$$\begin{cases} \overline{u} \cdot (x\overline{u} + y\overline{v} - \overline{b}) & = 0 \\ \overline{v} \cdot (x\overline{u} + y\overline{v} - \overline{b}) & = 0 \end{cases} \tag{1}$$

Två variabler, fortsättning

Vi kan skriva om ekvationssystemet (1) som

$$\begin{pmatrix} \overline{u} \cdot \overline{u} & \overline{u} \cdot \overline{v} \\ \overline{v} \cdot \overline{u} & \overline{v} \cdot \overline{v} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \overline{u} \cdot \overline{b} \\ \overline{v} \cdot \overline{b} \end{pmatrix}$$

vilket är samma sak som

$$\begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{2}$$

Normalekvationen

- Vi kan göra samma resonemang med m ekvationer och n variabler.
- Det ekvationssystem som motsvarar (2) kan alltid skrivas som

$$\boxed{A^t A\overline{x} = A^t\overline{b}.}$$

- Denna kallas *normalekvationen* för det överbestämda systemet $A\overline{x} = \overline{b}$.

Anpassa linje till mätdata

Om vi mäter upp $(x_1, x_2, \ldots, x_n)$ och $(y_1, y_2, \ldots, y_n)$ och förväntar oss att de ska ligga på en linje $y = kx + \ell$ får vi

$$\left\{\begin{array}{ccc} kx_1 + \ell & = & y_1 \\ kx_2 + \ell & = & y_2 \\ \vdots & \vdots & \vdots \\ kx_2 + \ell & = & y_2 \end{array}\right. \quad \Leftrightarrow \quad \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ \ell \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

och normalekvationen ges av

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ \ell \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Anpassa linje, fortsättning

Vi kan multiplicera ihop matriserna i normalekvationen och får

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} k \\ \ell \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

och det går i allmänhet att lösa ut k och ℓ explicit som

$$k = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}$$

och

$$\ell = \frac{\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i\right) - \left(\sum_{i=1}^n x_i y_i\right)\left(\sum_{i=1}^n x_i\right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}.$$