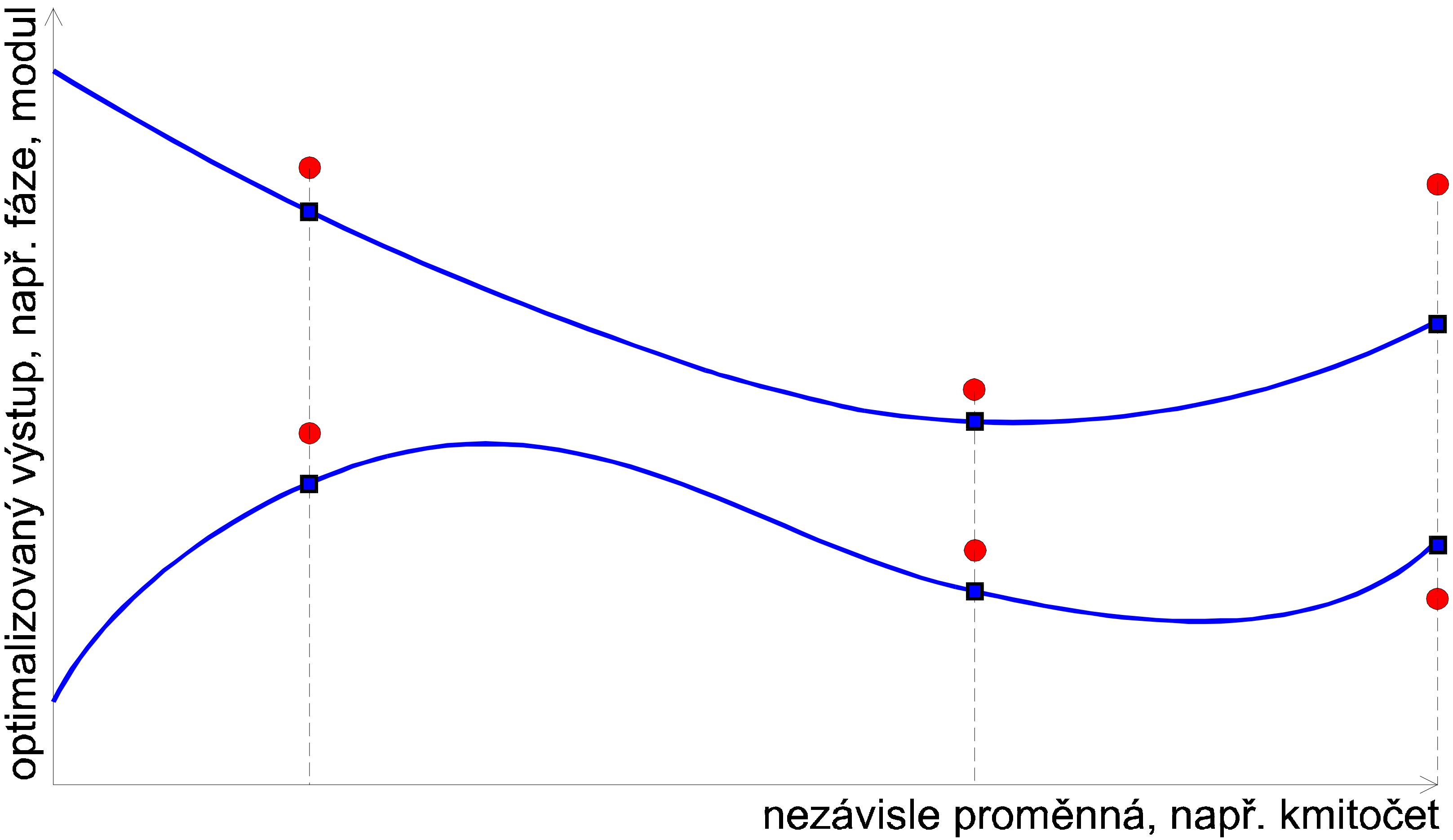
Optimalizace

Definice úlohy součtem nejmenších čtverců

Optimalizace

Definice úlohy součtem nejmenších čtverců



definice rozdílu mezi naměřenými a proloženými:

$$S(x_1,\ldots,x_n) = \sum_{k=1}^{m} R_k^2(x_1,\ldots,x_n)$$

$x_1,\ldots,x_n$ ... neznámé parametry obvodu

$R_k, k = 1,\ldots,m$ ... odchylky sledovaných výstupních obvodových veličin od uživatelem zadaných hodnot definice rozdílu mezi naměřenými a proloženými:

$$S(x_1,\ldots,x_n) = \sum_{k=1}^{m} R_k^2(x_1,\ldots,x_n)$$

$x_1,\ldots,x_n$ ... neznámé parametry obvodu

$R_k,\ k = 1,\ldots,m$ ... odchylky sledovaných výstupních obvodových veličin od uživatelem zadaných hodnot

minimum funkce více proměnných:

$$\operatorname{grad} S = \sum_{k=1}^{m} 2R_k \operatorname{grad} R_k = \mathbf{0}$$

Řešení úlohy zobecněnou metodou nejmenších čtverců

Řešení úlohy zobecněnou metodou nejmenších čtverců

odchylky $R_k^{(l)}$ v l-té iteraci optimalizace rozvineme pomocí MacLaurinovy řady

$$R_k^{(l)} \approx r_k + \left[\frac{\partial r_k}{\partial x_1}, \ldots, \frac{\partial r_k}{\partial x_n}\right] \begin{bmatrix} \Delta x_1^{(l)} \\ \vdots \\ \Delta x_n^{(l)} \end{bmatrix}$$

přičemž v této rovnici bylo použito stručnější označení

$$r_k = R_k\left(\boldsymbol{x}^{(l)}\right), \frac{\partial r_k}{\partial x_i} = \frac{\partial R_k}{\partial x_i}\left(\boldsymbol{x}^{(l)}\right)$$

po dosazení tohoto vztahu do úvodní rovnice dostaneme

$$\sum_{k=1}^{m}\left\{r_k+\left[\frac{\partial r_k}{\partial x_1},\ldots,\frac{\partial r_k}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots\\ \Delta x_n^{(l)}\end{bmatrix}\right\}\operatorname{grad}\left\{r_k+\left[\frac{\partial r_k}{\partial x_1},\ldots,\frac{\partial r_k}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots\\ \Delta x_n^{(l)}\end{bmatrix}\right\}\approx$$

$$\sum_{k=1}^{m}\left\{r_k+\left[\frac{\partial r_k}{\partial x_1},\ldots,\frac{\partial r_k}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots\\ \Delta x_n^{(l)}\end{bmatrix}\right\}\operatorname{grad} r_k=\mathbf{0}$$

zanedbáme-li i v této rovnici vliv druhých derivací rovnici přepíšeme do tvaru

$$\sum_{k=1}^{m}\left[\frac{\partial r_k}{\partial x_i}\frac{\partial r_k}{\partial x_1},\ldots,\frac{\partial r_k}{\partial x_i}\frac{\partial r_k}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots \\ \Delta x_n^{(l)}\end{bmatrix} = -\sum_{k=1}^{m}\frac{\partial r_k}{\partial x_i} r_k,\ i = 1,\ldots,n$$

rovnici přepíšeme do tvaru

$$\sum_{k=1}^{m}\left[\frac{\partial r_k}{\partial x_i}\frac{\partial r_k}{\partial x_1},\ldots,\frac{\partial r_k}{\partial x_i}\frac{\partial r_k}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots\\ \Delta x_n^{(l)}\end{bmatrix} = -\sum_{k=1}^{m}\frac{\partial r_k}{\partial x_i}r_k,\ i=1,\ldots,n$$

čili

$$\left[\frac{\partial r_1}{\partial x_i}\frac{\partial r_1}{\partial x_1}+\ldots+\frac{\partial r_m}{\partial x_i}\frac{\partial r_m}{\partial x_1},\ldots,\frac{\partial r_1}{\partial x_i}\frac{\partial r_1}{\partial x_n}+\ldots+\frac{\partial r_m}{\partial x_i}\frac{\partial r_m}{\partial x_n}\right]\begin{bmatrix}\Delta x_1^{(l)}\\ \vdots\\ \Delta x_n^{(l)}\end{bmatrix} = -\left[\frac{\partial r_1}{\partial x_i},\ldots,\frac{\partial r_m}{\partial x_i}\right]\begin{bmatrix}r_1\\ \vdots\\ r_m\end{bmatrix}$$

což lze vyjádřit přehledněji v maticové formě

$$\boldsymbol{J}^T\boldsymbol{J}\,\Delta\boldsymbol{x}^{(l)} = -\boldsymbol{J}^T\boldsymbol{r},\, l = 1,\ldots,l_{\max}$$

$$\boldsymbol{J} = \begin{bmatrix} \dfrac{\partial r_1}{\partial x_1} & \cdots & \dfrac{\partial r_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial r_m}{\partial x_1} & \cdots & \dfrac{\partial r_m}{\partial x_n} \end{bmatrix} \quad \Delta\boldsymbol{x}^{(l+1)} = \boldsymbol{x}^{(l)} + \Delta\boldsymbol{x}^{(l)}$$

což lze vyjádřit přehledněji v maticové formě

$$\boldsymbol{J}^T\boldsymbol{J}\,\Delta\boldsymbol{x}^{(l)} = -\boldsymbol{J}^T\boldsymbol{r},\ l = 1,\ldots,l_{\max}$$

$$\boldsymbol{J} = \begin{bmatrix} \dfrac{\partial r_1}{\partial x_1} & \cdots & \dfrac{\partial r_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial r_m}{\partial x_1} & \cdots & \dfrac{\partial r_m}{\partial x_n} \end{bmatrix} \quad \Delta\boldsymbol{x}^{(l+1)} = \boldsymbol{x}^{(l)} + \Delta\boldsymbol{x}^{(l)}$$

v praxi se tato metoda kombinuje s metodou gradientní

$$\Delta\boldsymbol{x}^{(l)} = -2\boldsymbol{J}^T\boldsymbol{r},\ l = 1,\ldots,l_{\max}$$

což lze vyjádřit přehledněji v maticové formě

$$\boldsymbol{J}^T \boldsymbol{J}\, \Delta \boldsymbol{x}^{(l)} = -\boldsymbol{J}^T \boldsymbol{r}, l = 1,\ldots,l_{\max}$$

$$\boldsymbol{J} = \begin{bmatrix} \dfrac{\partial r_1}{\partial x_1} & \cdots & \dfrac{\partial r_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial r_m}{\partial x_1} & \cdots & \dfrac{\partial r_m}{\partial x_n} \end{bmatrix} \quad \Delta \boldsymbol{x}^{(l+1)} = \boldsymbol{x}^{(l)} + \Delta \boldsymbol{x}^{(l)}$$

v praxi se tato metoda kombinuje s metodou gradientní

$$\Delta \boldsymbol{x}^{(l)} = -2 \boldsymbol{J}^T \boldsymbol{r}, l = 1,\ldots,l_{\max}$$

do finální Levenbergovy-Marquardtovy modifikace:

$$\left(\boldsymbol{J}^T \boldsymbol{J} + \lambda^{(l)}\, \mathbf{1}\right) \Delta \boldsymbol{x}^{(l)} = -\boldsymbol{J}^T \boldsymbol{r}, l = 1,\ldots,l_{\max}$$