Přednáška 6.

- Lineární modely
- BLUE (best linear unbiased estimators)

Základní lineární model

Mějme model náhodné veličiny

$$X_i = A + B \cdot i + w_i$$

- $i = 1, \dots, N$
- A, B jsou neznámé parametry, které chceme odhadovat
- $w_i \sim N(0, \sigma^2)$

Takový model nazveme **lineární** a lze jej zapsat ve tvaru

$$\boldsymbol{x} = \boldsymbol{H\theta} + \boldsymbol{w}$$

Kde $\boldsymbol{\theta} = [A, B]^T$ a $\boldsymbol{H} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & N \end{bmatrix}$ a $\boldsymbol{w} \sim N(0, \sigma^2 \boldsymbol{I})$.

Odhady parametrů základního lineárního modelu

Best linear unbiased estimator

Mějme model náhodné veličiny

$$\boldsymbol{x} = \boldsymbol{H\theta} + \boldsymbol{w}$$

Kde $\boldsymbol{x}$ je $N \times 1$ vektor pozorování, $\boldsymbol{\theta}$ je $p \times 1$ odhadovaný vektor parametrů , $\boldsymbol{H}$ je známá matice $N \times p\ (N > p)$, a $\mathbf{w}$ je $N \times 1$ vektor šumu, tedy $\mathbf{w} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$.

Pak odhad s minimálním rozptylem je

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{H}^T \boldsymbol{H}\right)^{-1} \boldsymbol{H}^T \boldsymbol{x}$$

a navíc kovarianční matice odhadu $\widehat{\boldsymbol{\theta}}$ je $cov\left(\widehat{\boldsymbol{\theta}}\right) = \sigma^2 \left(\boldsymbol{H}^T \boldsymbol{H}\right)^{-1}$.

Pro lineární modely se dokonce jedná o eficientní odhad (dosahuje CRLB) a je tedy nejlepší nestranný (best linear unbiased estimator = BLUE).

Obecný lineární model

Mějme model náhodné veličiny

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{\theta} + \boldsymbol{s} + \boldsymbol{w}$$

Kde $\boldsymbol{x}$ je $N \times 1$ vektor pozorování, $\boldsymbol{\theta}$ je $p \times 1$ odhadovaný vektor parametrů , $\boldsymbol{H}$ je známá $N \times p$ matice $(N > p)$, $\boldsymbol{s}$ je $N \times 1$ známý signál a $\mathbf{w}$ je $N \times 1$ vektor šumu, kde $\mathbf{w} \sim N(\mathbf{0}, \boldsymbol{C})$.

Pak odhad s minimálním rozptylem je

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{H}^T\boldsymbol{C}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{C}^{-1}(\boldsymbol{x} - \boldsymbol{s})$$

a navíc kovarianční matice odhadu $\widehat{\boldsymbol{\theta}}$ je $cov\left(\widehat{\boldsymbol{\theta}}\right) = \left(\boldsymbol{H}^T\boldsymbol{C}^{-1}\boldsymbol{H}\right)^{-1}$.

Pro lineární modely je odhad opět eficientní (dosahuje CRLB) a je tedy nejlepší nestranný.

Cvičení 6

1. Uvažujte pro zadaná data model náhodné veličiny dle lineárního modelu $x_k = A + B \cdot k + w_k$, $w_k \sim N(0, \sigma^2)$. Odhadněte A a B. Pomocí Monte Carlo simulace odhadněte rozptyly odhadů a porovnejte s teoretickým rozptylem.
   $[A_{true} = 3\ , B_{true} = 1]$
2. Stejnou úlohu řešte pro model $x_k = A + B \cdot k + C \cdot k^2 + w_k\ , w_k \sim N(0, \sigma^2)$. Odhadněte A, B a C.
   $[A_{true} = 3\ , B_{true} = 1, C_{true} = 0.1]$
3. Stejnou úlohu řešte pro model $x_k = A + B \cdot k + C \cdot k^2 + w_k$, kdy $w_k \sim N(0, 0.1k^2)$. Odhadněte A, B a C.
   $[A_{true} = 3\ , B_{true} = 1, C_{true} = 0.1]$
4. Uvažujte odhad parametrů lineárního modelu
   $$x[n] = \sum_{k=1}^{M} a_k \cos\frac{2\pi kn}{N} + \sum_{k=1}^{M} b_k \sin\frac{2\pi kn}{N} + w[n]$$
   kde $n = 0, \dots, N-1$ a $w[n] \sim N(0, \sigma^2)$.
5. Odvoďte vzorec pro odhad parametrů základního lineárního modelu pomocí nejmenších čtverců tj.
   $$\hat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}} \|\boldsymbol{x} - \boldsymbol{H}\boldsymbol{\theta}\|^2$$