přednáška 11

- Volba apriorní hustoty pravděpodobnosti
- MMSE vs. MAP
- Wienerův filtr

Volba apriorní hustoty

Připomenutí: aposteriorní hustota pravděpodobnosti má tvar

$$p(A|x) = \frac{p(x|A)p(A)}{p(x)} = \frac{p(x|A)p(A)}{\int p(x|A)p(A)dA}$$

Kde hustota $p(A)$ je apriorní hustota pravděpodobnosti.

Nejběžnější volby pro $p(A)$ jsou uniformní a normální pdf.

1. Pokud $p(x|A)$ modelujeme jako hustotu normálního rozdělení, pak díky volbě $A \sim U(a, b)$ nebo $A \sim N\left(\mu_A, \sigma_A^2\right)$ bude i aposteriorní hustota $p(A|x)$ normální.
2. Tomuto odpovídá např. model dat $x = A + w$, kde $w \sim N(\mu, \sigma^2)$. Pro odhady $\hat{A}$ se dá ukázat, že

   $a)$ $\hat{A} = \frac{\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma_A^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$, pokud $A \sim N\left(\mu_A, \sigma_A^2\right)$.

   b) Aposteriorní rozptyl je dán jako $var(A|x) = \sigma_{A|x}^2 = \frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$.

   $c)$ $BMSE\left(\hat{A}\right) < \frac{\sigma^2}{N}$

Metody bayesovských odhadů

Minimum Mean Square Error

$$Bmse(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right] = \int \left[\int (\hat{\theta} - \theta)^2 p(\theta|x) d\theta\right] p(x) dx$$

$$\hat{\theta} = E[\theta|x] = \int \theta p(\theta|x) d\theta$$

Maximum A posteriori Estimators

$$\hat{\theta} = \arg\max_{\theta} p(\theta|x)$$

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}$$

$$\hat{\theta} = \arg\max_{\theta} p(x|\theta)p(\theta)$$

Příklad: Odhad parametru θ exponenciálního rozdělení, pokud předpokládáme apriorní hustotu jako $p(\theta) = \lambda e^{-\lambda\theta}$, pro $\theta > 0$ a $p(\theta) = 0$ jinak.

Lineární bayesovský odhad

Linear minimum MSE (LMMSE):

$$\hat{\theta} = \sum_{k=1}^{N} a_k x_k + a_0$$

- Založeno na korelaci mezi odhadovaným parametrem θ a daty x_k, k=1,…,N. Parametr nekorelovaný s daty nemůže být takto odhadován.
- Není potřeba znát pdf $p(x, \theta)$.
- a_0 je potřeba pouze pokud x, θ mají nenulové střední hodnoty.
- Pouze suboptimální v případě nelineárních modelů.
- Koeficienty a_k nalezneme minimalizací $\mathrm{Bmse}(\hat{\theta}) = E\left[\left(\theta - \hat{\theta}\right)^2\right]$.

Příklad:

Odhadujeme θ pouze z jediného pozorování $x_1 \sim N(0, \sigma^2)$, skutečná hodnota je $\theta = x_1^2$, tudíž nejlepší odhad by byl $\hat{\theta} = x_1^2$.

LMMSE bude $\hat{\theta} = a_0 + a_1 x_1$, ale …

Lineární bayesovský odhad

Optimální volba koeficientů dává LMMSE ve tvaru:

$$\hat{\theta} = E[\theta] + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(x - E[x])$$

- $\boldsymbol{C}_{xx}$ je $N \times N$ kovarianční matice náhodné veličiny x
- $\boldsymbol{C}_{\theta x} = \boldsymbol{C}_{x\theta}^{T}$ je $1 \times N$ kovariance mezi θ a x

Vektorová verze:

- Koeficienty hledáme tak, aby minimalizovali $\mathrm{Bmse}(\hat{\theta}_i) = E\left[\left(\theta_i - \hat{\theta}_i\right)^2\right]$ pro všechny prvky vektorového parametru θ $(i = 1, \dots, d)$.
- LMMSE je tedy tvaru:

$$\hat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] + \mathbf{C}_{\boldsymbol{\theta} x}\mathbf{C}_{xx}^{-1}(x - E[x])$$

- Kde $\boldsymbol{C}_{xx}$ je $N \times N$ kovarianční matice náhodné veličiny x
- $\boldsymbol{C}_{\theta x} = \boldsymbol{C}_{x\theta}^{T}$ je $\mathrm{d} \times N$ kovariance mezi θ a x

Signal processing – Wiener filter

Uvažujme $\boldsymbol{x} = [x[1], x[2], \dots, x[N]]$ náhodný vektor s nulovou střední hodnotou a kovariancí

$$\boldsymbol{C}_{xx} = \begin{bmatrix} r_{xx}[1] & \cdots & r_{xx}[N] \\ \vdots & \ddots & \vdots \\ r_{xx}[N] & \cdots & r_{xx}[1] \end{bmatrix}$$

Tedy r_k je ACF (autocorrelation function) procesu $x[N+1]$ a $\boldsymbol{C}_{xx}$ je jeho autokorelační matice.

Wienerovými filtry nazveme 3 základní aplikace toho problému:

1. Filtrování = odhadujeme signál $\theta = s[N]$ v modelu $x[k] = s[k] + \mathrm{w}[k], k = 1, \dots N$ (pro odhad používáme pouze data přítomná a minulá (pro odhad $s[2]$ mohu využít jen $x[1], x[2]$)
2. Vyhlazování (smoothing) = stejné jako filtrování, ale nad celými daty.
3. Predikce = odhadujeme $\theta = s[N+k]$ na základě $x[1], x[2], \dots, x[N]$

Ve všech případech uvažujeme $E[\boldsymbol{\theta}] = E[\boldsymbol{x}] = 0$, tedy

$$\widehat{\boldsymbol{\theta}} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\boldsymbol{x}$$

cvičení 11

- Házejte mincí a odhadujte z dat pravděpodobnost, že padne orel.
- Majitel e-shopu má data o denních prodejích za prvních 290 dní v roce. Přijde mu, že za posledních 16 dní jdou prodeje nahoru a chtěl by na základě této hypotézy odhadnout, kolik bude mít prodejů do konce roku.