přednáška 11

- Opakování
- Lineární bayesovský MMSE odhad (LMMSE)
- 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) $\mathrm{BMSE}(\hat{\mathrm{A}}) < \frac{\sigma^2}{\mathrm{N}}$

Metody bayesovských odhadů

Minimum Mean Square Error (MMSE)

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

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

Maximum A posteriori Estimators

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

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

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

Často je obtížné odvodit tyto odhady analyticky. MMSE vyžaduje odvození celé aposteriorní hustoty, což obnáší integraci $p(x|\theta)p(\theta)$. MAP vyžaduje hledání globálního maxima mnohorozměrné funkce.

Lineární bayesovský odhad

Zjednodušení: Uvažujeme lineární (skalární) bayesovský MMSE odhad (LMMSE):

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

- θ je náhodný parametr, $\boldsymbol{x}$ a θ mají sdružené pdf $p(\boldsymbol{x}, \theta)$, nebude ho ale potřeba znát.
- Ukáže se, že je třeba předpokládat korelaci mezi θ a $\boldsymbol{x}$. Parametr nekorelovaný s daty nemůže být takto odhadován.
- a_0 je potřeba pouze pokud $\boldsymbol{x}$ a θ mají nenulové střední hodnoty.
- Odhad bude suboptimální v případě negaussovských modelů.
- Koeficienty a_k nalezneme minimalizací $\mathrm{Bmse}(\hat{\theta}) = E\left[\left(\theta - \hat{\theta}\right)^2\right]$.

Příklad (co se stane, když parametr není korelovaný s daty):

Odhadujeme θ pouze z jediného pozorování $x_1 \sim N(0, \sigma^2)$, kdy $\theta = x_1^2$, tudíž nejlepší odhad by byl $\hat{\theta} = x_1^2$. LMMSE je $\hat{\theta} = a_0 + a_1 x_1$. Dořešte příklad.

Lineární bayesovský odhad

Optimální volba koeficientů dává LMMSE ve tvaru:

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

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

Vektorový parametr $\boldsymbol{\theta}$:

- 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 $\boldsymbol{\theta}$ $(i = 1, \dots, d)$.
- LMMSE je tedy tvaru:

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

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

Vidíme, že pro výpočet odhadu nám stačí střední hodnoty a kovariance $\boldsymbol{\theta}$ a $\boldsymbol{x}$, pdf může být libovolné.

Wienerovy filtry ve zpracování signálů

Lineární bayesovské odhady jsou obecně nazývány **Wienerovy filtry.**

Uvažujme $\boldsymbol{x} = \left[x[1], x[2], \dots, x[N]\right]$ náhodný vektor s nulovou střední hodnotou a kovariancí $\boldsymbol{C}_{\boldsymbol{xx}}$. Pokud je $\boldsymbol{x}$ tzv. slabě stacionární proces – WSS (zjednodušeně: střední hodnota a kovariance jsou nezávislé na čase), pak

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

kde $r_{xx}[t] = \mathrm{E}[x[n]x[n+t]]$ je autokorelační funkce ($\boldsymbol{x}$ je WSS takže $r_{xx}[t]$ nezávisí na n).

Tři základní aplikace LMMSE ve zpracování signálů (Wienerův filtr):

1. **Odšumování (on-line denoising):** odhadujeme $\theta = s[n]$ v modelu $x[m] = s[m] + \mathrm{w}[m], m = 1, \dots, n$ (pro odhad používáme pouze data přítomná a minulá (např. pro odhad $s[2]$ mohu využít jen $x[1], x[2]$). Hledáme kauzální filtr.
2. **Vyhlazování (off-line smoothing):** $\theta = s[n]$ má být odhadnuto pro $n = 1, \dots, N$ za použití celé sady měření $\boldsymbol{x} = \left[x[1], x[2], \dots, x[N]\right]$ (např. pro odhad $s[2]$ máme k dispozici celé $\boldsymbol{x}$).
3. **Predikce:** odhadujeme $\theta = s[N+k]$, $k > 0$, za použití $\boldsymbol{x} = \left[x[1], x[2], \dots, x[N]\right]$.

Otázky na tělo: Kde je ten „filtr“? Může se Wienerův filtr měnit v čase?

Cvičení 11

- 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.