

přednáška 10

- ▶ Volba apriorní hustoty pravděpodobnosti
- ▶ MMSE vs. MAP

Volba apriorní hustoty

Připomenutí: aposteriorní hustota pravděpodobnosti má tvar

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

Kde hustota $p(\theta)$ je apriorní hustota pravděpodobnosti.

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

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

$$a) \hat{\theta}_{\text{MSE}} = \frac{\frac{N}{\sigma^2}\bar{x} + \frac{\mu_\theta}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}, \text{ pokud } \theta \sim N(\mu_\theta, \sigma_\theta^2).$$

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

$$c) \text{ BMSE}(\hat{\theta}) < \frac{\sigma^2}{N}$$

Metody bayesovských odhadů

Minimum Mean Square Error

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

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

Maximum A posteriori Estimators

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

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

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

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.

cvičení 10

- ▶ Házejte mincí a odhadujte z dat pravděpodobnost, že padne orel. Použijte apriorní předpoklad, abyste se vyvarovali špatným závěrům z malého počtu pozorování (když např. dvakrát padne orel, mohlo by se usoudit, že pravděpodobnost je 100%).
- ▶ Majitel e-shopu má data o denních prodejkách za prvních 290 dní v roce (data na e-learningu). 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.