Recap: Inference for Gaussian parameters

Devdatt Dubhashi

A sensor problem

- Noisy measurements from a sensor: $x_1, x_2, \cdots, x_n$
- Where is it located?

A simple Gaussian Model

- Data modelled and generated i.i.d. from a Gaussian Distribution

$$\mathcal{N}(\mu, \sigma^2)$$

- Where μ is the true but unknown position of the sensor and the variance σ^2 is a noise parameter which is known.

Likelihood

$$P(x_1, \cdots, x_n \mid \mu, \sigma^2) = \prod_{i}^{n} \mathcal{N}(x_i \mid \mu, \sigma^2)$$

Prior

$$P(\mu) = \mathcal{N}(\mu \mid \mu_0, \sigma_0^2)$$

μ_0 and σ_0^2 are hyperparameters that are known from prior domain knowledge

Posterior

- Bayes Rule:

likelihood

prior

$$P(\mu \mid x_1, \ldots, x_n) = \frac{\mathcal{N}(x_1, \cdots, x_n \mid \mu, \sigma^2)\mathcal{N}(\mu \mid \mu_0, \sigma_0^2)}{\int_{-\infty}^{+\infty} \mathcal{N}(x_1, \cdots, x_n \mid \mu, \sigma^2)\mathcal{N}(\mu \mid \mu_0, \sigma_0^2) d\mu}$$

Normalizing
factor: Bad Shit!

Conjugate Prior!

- Since we have used a conjugate prior to the likelihood, we know immediately that the resulting posterior distribution is also a Gaussian distribution without having to compute the denominator.

$$\mathcal{N}(\mu \mid \mu_n, \sigma_n)$$

- What are the parameters μ_n, σ_n ?
- Pattern match with numerator of Bayes formula!

Pattern Match!

$$P(\mu \mid x_1, \ldots, x_n) = \frac{\mathcal{N}(x_1, \cdots, x_n \mid \mu, \sigma^2)\mathcal{N}(\mu \mid \mu_0, \sigma_0^2)}{\int_{-\infty}^{+\infty} \mathcal{N}(x_1, \cdots, x_n \mid \mu, \sigma^2)\mathcal{N}(\mu \mid \mu_0, \sigma_0^2) d\mu}$$

$$\mathcal{N}(\mu \mid \mu_n, \sigma_n)$$

Match exponents

$$\exp\left(-\frac{1}{2\sigma_n^2}(\mu-\mu_n)^2\right)$$

$$\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right)$$

Pattern Match

- Compare coefficient of μ^2:

$$\frac{1}{\sigma_n^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}$$

- Compare coefficient of μ and use above:

$$\mu_n = \frac{\sigma^2}{n\sigma_0^2 + \sigma^2}\mu_0 + \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2}\mu_{ML}$$

$$\mu_{ML} = \frac{\sum_i x_i}{n}$$

Model Selection: Which model is better?

- Suppose we have two models with variances

$$\sigma_1^2 \quad \text{and} \quad \sigma_2^2$$

- Which should one prefer?
- Compute Bayes Factor:

$$\frac{P_{M_1}(x_1, \cdots x_n)}{P_{M_2}(x_1, \cdots x_n)} = \frac{\mathcal{N}(x_1, \cdots, x_n \mid \mu_{n,1}, \sigma_{n,1}^2)}{\mathcal{N}(x_1, \cdots, x_n \mid \mu_{n,2}, \sigma_{n,2}^2)}$$