Gaussian Facts

For a multivariate Gaussian $p(x_a, x_b)$ with variables partitioned into $x_a$ and $x_b$, we have:

$p(x_a)$ is also Gaussian

and

$p(x_a | x_b)$ is also Gaussian

Sampling Continuous Distributions

- Suppose you want to generate samples from (i.e., simulate a probability distribution).
- The typical tool at your disposal is a pseudo random number generator returning approximately uniformly distributed rational numbers in $[0,1]$
- Sampling Bernoulli processes is straightforward
- Variants of uniform distributions are also easy
- Example: $p(x) = \begin{cases} 5 & x \in [0.4, 0.6] \\ 0 & \text{otherwise} \end{cases}$

Sampling Continuous Distributions

- $N(0,1)$ is less obvious (there are standard fast methods)
- A general approach for sampling a continuous distribution (sometimes call inverse transformation sampling) is based on the cumulative distribution function, CDF, denoted by $F(x)$

Cumulative Distribution Function

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(x) \, dx \quad \text{(continuous distributions)}$$
Sampling Continuous Distributions

• To sample a distribution $p(x)$

  Prepare approximation of $F^{-1}(x)$

  Loop
  sample $x \in [0,1]$
  report $F^{-1}(x)$

Example (from Bishop, PRML)
Estimating the mean of a univariate Gaussian

Assume that the variance is known.
Given data points $x_i$, what is the "best" estimate for the mean?

The maximum likelihood estimate is $\mu_{ML} = \frac{1}{N} \sum x_i$

But what if the number of points is small?

Let's consider the case where we want to incorporate prior information.

IE, let's do Bayes.

Example (from Bishop, PRML)
Estimating the mean of a univariate Gaussian

$$p(\mu \mid \{x_i\}) \propto p(\mu) p(\{x_i\} \mid \mu)$$

$$= p(\mu) \prod_i p(\{x_i\} \mid \mu)$$

$$\propto p(\mu) \prod_i \exp\left(-\left(x_i - \mu\right)^2\right)$$

What should we use for $p(\mu)$?

Example (from Bishop, PRML)
Estimating the mean of a univariate Gaussian

$$p(\mu \mid \{x_i\}) \propto p(\mu) \prod_i \exp\left(-\left(x_i - \mu\right)^2\right)$$

By inspection, if $p(\mu) \propto \exp\left(-\left(\mu_0 - \mu\right)^2\right)$ then
the form of the posterior is the same as the prior.

IE, given known variance, a conjugate prior for the mean of the Gaussian is a Gaussian.

We use a different mean for the prior.
Example (from Bishop, PRML)

### Estimating the mean of a univariate Gaussian

\[
p(\mu \mid \{x_i\}) \propto \exp \left( \frac{- (\mu_0 - \mu)^2}{\sigma_0^2} \right) \prod_i \exp \left( \frac{- (x_i - \mu)^2}{\sigma^2} \right)
\]

To find the MAP (maximum a posteriori) estimate we maximize.

Maximizing is the same as minimizing the negative log.

\[
-\log(p(\mu \mid \{x_i\})) = \frac{(\mu_0 - \mu)^2}{\sigma_0^2} + \frac{N\mu}{\sigma^2} + \frac{N}{\sigma^2} \sum_i x_i = \mu_0 \frac{1}{\sigma_0^2} + \mu \frac{N}{\sigma^2} + \frac{1}{\sigma^2} \sum_i x_i
\]

differentiating and setting derivatives to zero gives

\[
\mu = \frac{\mu_0 + \frac{N}{\sigma^2} \mu_{ML}}{\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}} = \frac{\mu_0 \frac{1}{\sigma_0^2} + \mu \frac{N}{\sigma^2} + \frac{1}{\sigma^2} \sum_i x_i}{\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}} = \frac{\frac{1}{\sigma_0^2} \mu_0 + \frac{N}{\sigma^2} \mu_{ML}}{\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}} + \frac{N \sigma_0^2}{\sigma^2 + N \sigma_0^2} \mu_{ML}
\]

Example (from Bishop, PRML)

### Unknown variance or mean and variance

Similar stories can be told if the mean is known and the variance is not, or both are unknown. We will only consider the conjugate priors.

We will simplify things by using the inverse of the covariance matrix which is called the precision matrix.

In the univariate case this is simply: \( \lambda = \frac{1}{\sigma^2} \)

Example (from Bishop, PRML)

### Estimating the variance

\[
p(\{x_i\} \mid \lambda) = \prod_i N(x_i \mid \mu, \lambda) \propto \lambda^{N/2} \exp \left\{ -\frac{\lambda}{2} \sum_i (x_i - \mu)^2 \right\}
\]

Inspection reveals that multiplying this by a gamma distribution

\[
\text{Gam}(\lambda \mid a, b) = \frac{1}{\Gamma(a)} b^a \lambda^{a-1} \exp(-b \lambda)
\]

yields a posterior of the same form.