Laplace’s Method

The idea behind the Laplace approximation is simple. We assume that an unnormalized probability density $P(x)$, whose normalizing constant

$$Z_P \equiv \int P^*(x) \, dx$$

is of interest, has a peak at a point $x_0$. We Taylor-expand the logarithm of $P^*(x)$ around this peak:

$$\ln P^*(x) \simeq \ln P^*(x_0) - \frac{c}{2}(x - x_0)^2 + \cdots,$$

where

$$c = -\frac{\partial^2}{\partial x^2} \ln P^*(x) \bigg|_{x=x_0}. \quad (27.3)$$

We then approximate $P^*(x)$ by an unnormalized Gaussian,

$$Q^*(x) \equiv P^*(x_0) \exp \left[ -\frac{c}{2}(x - x_0)^2 \right], \quad (27.4)$$

and we approximate the normalizing constant $Z_P$ by the normalizing constant of this Gaussian,

$$Z_Q = P^*(x_0) \sqrt{\frac{2\pi}{c}}. \quad (27.5)$$

We can generalize this integral to approximate $Z_P$ for a density $P^*(x)$ over a $K$-dimensional space $x$. If the matrix of second derivatives of $-\ln P^*(x)$ at the maximum $x_0$ is $A$, defined by:

$$A_{ij} = -\frac{\partial^2}{\partial x_i \partial x_j} \ln P^*(x) \bigg|_{x=x_0}, \quad (27.6)$$

so that the expansion (27.2) is generalized to

$$\ln P^*(x) \simeq \ln P^*(x_0) - \frac{1}{2}(x - x_0)^T A (x - x_0) + \cdots, \quad (27.7)$$

then the normalizing constant can be approximated by:

$$Z_P \simeq Z_Q = P^*(x_0) \frac{1}{\sqrt{\det \frac{1}{2\pi} A}} = P^*(x_0) \sqrt{\frac{(2\pi)^K}{\det A}}. \quad (27.8)$$

Predictions can be made using the approximation $Q$. Physicists also call this widely-used approximation the saddle-point approximation.
The fact that the normalizing constant of a Gaussian is given by
\[ Z = \sqrt{\frac{(2\pi)^k}{\det A}} \] (27.9)
can be proved by making an orthogonal transformation into the basis \( u \) in which \( A \) is transformed into a diagonal matrix. The integral then separates into a product of one-dimensional integrals, each of the form
\[ \int du_i \exp \left[ -\frac{1}{2} \lambda_i u_i^2 \right] = \sqrt{\frac{2\pi}{\lambda_i}}. \] (27.10)

The product of the eigenvalues \( \lambda_i \) is the determinant of \( A \).

The Laplace approximation is basis-dependent: if \( x \) is transformed to a nonlinear function \( u(x) \) and the density is transformed to \( P(u) = P(x) |dx/du| \) then in general the approximate normalizing constants \( Z_Q \) will be different. This can be viewed as a defect – since the true value \( Z_P \) is basis-independent – or an opportunity – because we can hunt for a choice of basis in which the Laplace approximation is most accurate.

### 27.1 Exercises

**Exercise 27.1.** (See also exercise 22.8 (p.307).) A photon counter is pointed at a remote star for one minute, in order to infer the rate of photons arriving at the counter per minute, \( \lambda \). Assuming the number of photons collected \( r \) has a Poisson distribution with mean \( \lambda \),
\[ P(r | \lambda) = \exp(-\lambda) \frac{\lambda^r}{r!}. \] (27.11)
and assuming the improper prior \( P(\lambda) = 1/\lambda \), make Laplace approximations to the posterior distribution
(a) over \( \lambda \)
(b) over \( \log \lambda \). [Note the improper prior transforms to \( P(\log \lambda) = \) constant.]

**Exercise 27.2.** Use Laplace’s method to approximate the integral
\[ Z(u_1, u_2) = \int_{-\infty}^{\infty} da f(a)^{u_1} (1 - f(a))^{u_2}, \] (27.12)
where \( f(a) = 1/(1 + e^{-a}) \) and \( u_1, u_2 \) are positive. Check the accuracy of the approximation against the exact answer (23.29, p.316) for \( (u_1, u_2) = (1/2, 1/2) \) and \( (u_1, u_2) = (1, 1) \). Measure the error (\( \log Z_P - \log Z_Q \)) in bits.

**Exercise 27.3.** Linear regression. \( N \) datapoints \( \{(x^{(n)}, t^{(n)})\} \) are generated by the experimenter choosing each \( x^{(n)} \), then the world delivering a noisy version of the linear function
\[ y(x) = w_0 + w_1x, \] (27.13)
\[ t^{(n)} \sim \text{Normal}(y(x^{(n)}), \sigma^2_n). \] (27.14)
Assuming Gaussian priors on \( w_0 \) and \( w_1 \), make the Laplace approximation to the posterior distribution of \( w_0 \) and \( w_1 \) (which is exact, in fact) and obtain the predictive distribution for the next datapoint \( t^{(N+1)} \), given \( x^{(N+1)} \).

(See MacKay (1992a) for further reading.)