Machine Learning and Bayesian Inference
How to evaluate Gaussian integrals

Sean B. Holden

1 Introduction

The following notes show how to evaluate the standard integral required in deriving the approximation to the Bayes-optimal neural network.

2 Gaussian integrals: the simple case

The simplest version of the problem is to evaluate the integral

$$I = \int_{-\infty}^{\infty} \exp\left(-\frac{ax^2}{2}\right)\, dx.$$

This is a fairly standard integration problem and several solutions are available in text books. For example, start by squaring it, so

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \exp\left(-\frac{ax^2}{2}\right)\, dx \times \int_{-\infty}^{\infty} \exp\left(-\frac{ay^2}{2}\right)\, dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{a}{2}(x^2+y^2)\right)\, dx\, dy. \end{aligned}$$

Then convert to polar co-ordinates, so $x = r\cos\theta$, $y = r\sin\theta$ and the Jacobian is

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r.$$

We now have

$$I^2 = \int_0^{2\pi}\int_0^{\infty} r\exp\left(-\frac{ar^2}{2}\right)\, dr\, d\theta$$

and as

$$-\frac{1}{a}\frac{d}{dr}\left(\exp(-\frac{ar^2}{2})\right) = r\exp(-\frac{ar^2}{2})$$

this is

$$I^2 = \int_0^{2\pi}\left[-\frac{1}{a}\exp(-\frac{ar^2}{2})\right]_0^{\infty} d\theta = \frac{1}{a}\int_0^{2\pi} d\theta = \frac{2\pi}{a}$$

and so

$$I = \sqrt{\frac{2\pi}{a}}.$$

3 Gaussian integrals: the general case

The problem now is to evaluate the more general integral

$$I = \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\left(\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c\right)\right) d\mathbf{x}$$

where $\mathbf{A}$ is an $n \times n$ symmetric matrix with real-valued elements, $\mathbf{b} \in \mathbb{R}^n$ is a real-valued vector and $c \in \mathbb{R}$. First of all, we can dispose of the constant part of the integrand as

$$I = \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\left(\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x}\right)\right) \exp\left(-\frac{c}{2}\right) d\mathbf{x} = \exp\left(-\frac{c}{2}\right) I'$$

where

$$I' = \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\left(\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x}\right)\right) d\mathbf{x}.$$

We're now going to make a change of variables, based on the fact that $\mathbf{A}$ has n eigenvalues v_i and n eigenvectors $\mathbf{e}_i$ such that

$$\mathbf{A}\mathbf{e}_i = v_i\mathbf{e}_i \tag{1}$$

for $i = 1, \ldots, n$. The eigenvalues can be found such that they are orthonormal

$$\mathbf{e}_i^T\mathbf{e}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Multiplying (1) on both sides by $\mathbf{A}^{-1}$ gives

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{e}_i = \mathbf{I}_n\mathbf{e}_i = \mathbf{e}_i = \mathbf{A}^{-1}v_i\mathbf{e}_i$$

for $i = 1, \ldots, n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix. Consequently

$$\mathbf{A}^{-1}\mathbf{e}_i = \frac{1}{v_i}\mathbf{e}_i$$

for $i = 1, \ldots, n$ and $\mathbf{A}^{-1}$ has the same eigenvectors as $\mathbf{A}$, but eigenvalues $1/v_i$. As the eigenvectors are orthonormal, any vector $\mathbf{x}$ can be written as

$$\mathbf{x} = \sum_{i=1}^{n} \lambda_i\mathbf{e}_i$$

for suitable values λ_i, and we can represent $\mathbf{b}$ as

$$\mathbf{b} = \sum_{i=1}^{n} \beta_i\mathbf{e}_i$$

in the same way. Next, we make a change of variables from $\mathbf{x}$ to

$$\boldsymbol{\lambda}^T = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \end{bmatrix}.$$

To make a change of variables we need to compute the Jacobian and rewrite the integral. The Jacobian for this transformation is

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial \lambda_1} & \frac{\partial x_2}{\partial \lambda_1} & \cdots & \frac{\partial x_n}{\partial \lambda_1} \\ \frac{\partial x_1}{\partial \lambda_2} & \frac{\partial x_2}{\partial \lambda_2} & \cdots & \frac{\partial x_n}{\partial \lambda_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial \lambda_n} & \frac{\partial x_2}{\partial \lambda_n} & \cdots & \frac{\partial x_n}{\partial \lambda_n} \end{vmatrix}.$$

As we saw above that

$$\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{e}_i$$

we have

$$x_j = \sum_{i=1}^{n} \lambda_i \mathbf{e}_i^{(j)}$$

where $\mathbf{e}_i^{(j)}$ is the jth element of $\mathbf{e}_i$, and so

$$\frac{\partial x_j}{\partial \lambda_k} = \mathbf{e}_k^{(j)}.$$

Thus

$$J = \begin{vmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \\ \vdots & \vdots & \cdots & \vdots \end{vmatrix}.$$

That is, the determinant of the matrix having the eigenvectors as its columns. Define

$$\mathbf{E} = \begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix}$$

such that $J = |\mathbf{E}|$. As the eigenvectors are orthonormal we have

$$J^2 = |\mathbf{E}||\mathbf{E}| = |\mathbf{E}||\mathbf{E}^T| = |\mathbf{E}\mathbf{E}^T| = |\mathbf{I}_n| = 1$$

and so $J = 1$.

Let's now look at the integrand

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x}.$$

Looking at the first term

$$\begin{aligned} \left(\sum_{i=1}^{n} \lambda_i \mathbf{e}_i^T \right) \mathbf{A} \left(\sum_{i=1}^{n} \lambda_i \mathbf{e}_i \right) &= \left(\sum_{i=1}^{n} \lambda_i \mathbf{e}_i^T \right) \left(\sum_{i=1}^{n} \lambda_i \mathbf{A} \mathbf{e}_i \right) \\ &= \left(\sum_{i=1}^{n} \lambda_i \mathbf{e}_i^T \right) \left(\sum_{i=1}^{n} \lambda_i v_i \mathbf{e}_i \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j v_i \mathbf{e}_j^T \mathbf{e}_i \\ &= \sum_{i=1}^{n} v_i \lambda_i^2. \end{aligned}$$

The second term simplifies in a similar way

$$\begin{aligned}
\mathbf{b}^T\mathbf{x} &= \left(\sum_{i=1}^{n} \beta_i \mathbf{e}_i^T\right) \left(\sum_{j=1}^{n} \lambda_j \mathbf{e}_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i \lambda_j \mathbf{e}_i^T \mathbf{e}_j \\
&= \sum_{i=1}^{n} \beta_i \lambda_i
\end{aligned}$$

and so the integrand becomes

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} = \sum_{i=1}^{n} \left(v_i\lambda_i^2 + \beta\lambda_i\right).$$

Thus the result of changing the variable is that

$$\begin{aligned}
I' &= \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\left(\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x}\right)\right) d\mathbf{x} \\
&= \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n} \left(v_i\lambda_i^2 + \beta\lambda_i\right)\right)\right) d\boldsymbol{\lambda} \\
&= \prod_{i=1}^{n} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(v_i\lambda_i^2 + \beta_i\lambda_i\right)\right) d\lambda_i.
\end{aligned}$$

What have we gained by changing the variable?

- We have changed a multiple integral into a *product of single integrals.*
- Each of these single integrals is *almost* of a form that can be solved using the simple case above.

How do we proceed? Writing

$$\left(-\frac{1}{2}\left(v_i\lambda_i^2 + \beta_i\lambda_i\right)\right) = -\frac{v_i}{2}\left(\lambda_i + \frac{\beta_i}{2v_i}\right)^2 + \frac{\beta_i^2}{8v_i}$$

and changing the variable in the simple integral from λ_i to

$$\theta_i = \left(\lambda_i + \frac{\beta_i}{2v_i}\right)$$

gives

$$\frac{d\theta_i}{d\lambda_i} = 1$$

and

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(v_i\lambda_i^2 + \beta_i\lambda_i\right)\right) d\lambda_i = \exp\left(\frac{\beta_i^2}{8v_i}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{v_i}{2}\theta_i^2\right) d\theta_i$$
$$= \exp\left(\frac{\beta_i^2}{8v_i}\right)\left(\frac{2\pi}{v_i}\right)^{1/2}$$

using the simple case. We now have

$$I' = \prod_{i=1}^{n} \exp\left(\frac{\beta_i^2}{8v_i}\right)\left(\frac{2\pi}{v_i}\right)^{1/2}.$$

This can be simplified further in two steps. First, if $\mathbf{A}$ has eigenvalues v_i then

$$|\mathbf{A}| = \prod_{i=1}^{n} v_i$$

and so

$$\prod_{i=1}^{n}\left(\frac{1}{v_i}\right) = |\mathbf{A}|^{-1}.$$

Thus

$$\prod_{i=1}^{n}\left(\frac{2\pi}{v_i}\right)^{1/2} = (2\pi)^{n/2}|\mathbf{A}|^{-1/2}.$$

Then, we have

$$\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} = \left(\sum_{i=1}^{n}\beta_i\mathbf{e}_i^T\right)\mathbf{A}^{-1}\left(\sum_{i=1}^{n}\beta_i\mathbf{e}_i\right)$$
$$= \left(\sum_{i=1}^{n}\beta_i\mathbf{e}_i^T\right)\left(\sum_{i=1}^{n}\frac{\beta_i}{v_i}\mathbf{e}_i\right)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\beta_j\mathbf{e}_j^T\mathbf{e}_i\frac{\beta_i}{v_i}$$
$$= \sum_{i=1}^{n}\frac{\beta_i^2}{v_i}.$$

Thus

$$\prod_{i=1}^{n}\exp\left(\frac{\beta_i^2}{8v_i}\right) = \exp\left(\frac{1}{8}\sum_{i=1}^{n}\frac{\beta_i^2}{v_i}\right)$$
$$= \exp\left(\frac{1}{8}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}\right)$$

and collecting everything together we have

$$
\begin{aligned}
I &= \exp\left(-\frac{c}{2}\right)(2\pi)^{n/2}|\mathbf{A}|^{-1/2}\exp\left(\frac{1}{8}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}\right) \\
&= (2\pi)^{n/2}|\mathbf{A}|^{-1/2}\exp\left(-\frac{1}{2}\left(c - \frac{\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}}{4}\right)\right).
\end{aligned}
$$