Comments

- I want to emphasize that parts of this documentation are not "clean" in the sense that some variable names might be used twice or it still contains some typos and shortcuts. If you don't understand part of it don't hesitate to ask me.
- I added a Background section where you can find a derivation for the Laplace approximation a definition of an exponential family and the application of a "trick" that becomes necessary for the sqrt transformation.

Background

0.1 Laplace Approximation

The Laplace approximation (LPA) is a tool to fit a normal distribution to the PDF of a given other distribution. The only constraints for the other distribution are: one peak (mode/ point of maximum) and twice differentiable. Laplace proposed a simple 2-term Taylor expansion on the log pdf. If $\hat{\theta}$ denotes the mode of a pdf $h(\theta)$, then it is also the mode of the log-pdf $q(\theta) = \log h(\theta)$. The 2-term Taylor expansion of $q(\theta)$ therefore is:

$$
\begin{aligned}
q(\theta) &\approx q(\hat{\theta}) + q'(\hat{\theta})(\theta - \hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})q''(\hat{\theta})(\theta - \hat{\theta}) && (1)\\
&= q(\hat{\theta}) + 0 + \frac{1}{2}(\theta - \hat{\theta})q''(\hat{\theta})(\theta - \hat{\theta}) \qquad [\text{since } q'(\theta) = 0] && (2)\\
&= c - \frac{(\theta - \mu)^2}{2\sigma^2} && (3)
\end{aligned}
$$

where c is a constant, $\mu = \hat{\theta}$ and $\sigma^2 = \{-q''(\hat{\theta})\}^{-1}$. The right hand side of the last line matches the log-pdf of a normal distribution $N(\mu, \sigma^2)$. Therefore the pdf $h(\theta)$ is approximated by the pdf of the normal distribution $N(\mu, \sigma^2)$ where $\mu = \hat{\theta}$ and $\sigma^2 = \{-q''(\hat{\theta})\}^{-1}$. Note, that even though this derivation is done for the one dimensional case only, it is also true for the multidimensional case. The second derivative just becomes the Hessian of the pdf at the mode.

0.2 Exponential Family

exponential family form

$$
f(x|\theta) = h(x)\exp[\eta(\theta)t(x) - A(\theta)] \tag{4}
$$

for sufficient statistics $t : X \to \mathbb{R}$, natural parameters $\eta : \Theta \to \mathbb{R}$, and functions $A : \Theta \to \mathbb{R}$ and $h : X \to \mathbb{R}_+$

0.3 Chi2 <-> Normal

It is already well-known that the Chi-squared distribution describes the sum of independent, standard normal random variables. To introduce a certain 'trick' we show the forward and backward transformation between chi2 and normal.
Let X be normal with $\mu = 0, \sigma^2 = 1$. Let $Y = X^2$ and therefore $g(x) = x^2$, which is neither monotone nor injective. Take $I_1 = (-\infty, 0)$ and $I_2 = [0, +\infty)$. Then g is monotone and injective on I_1 and I_2 and $I_1 \cup I_2 = \mathbb{R}$. $g(I_1) = (0, \infty)$ and $g(I_2) = [0, \infty)$. Then $g_1^{-1} : [0, \infty) \to \mathbb{R}$ by $g_1^{-1}(y) = -\sqrt{y}$ and $g_2^{-1} : [0, \infty) \to \mathbb{R}$ by $g_2^{-1}(y) = \sqrt{y}$. Then

$$
\left|\frac{\partial g_i^{-1}(y)}{\partial y}\right| = \left|\frac{1}{2\sqrt{y}}\right| = \frac{1}{2\sqrt{y}}
$$

Applying Equation **??** we can transform a normal distribution to a chi-squared.

$$
\begin{aligned}
f_Y(y) &= f_X(g_1^{-1}(y)) \left| \frac{\partial g_1^{-1}(y)}{\partial y} \right| \mathbf{1}_\wedge(y) + f_X(g_2^{-1}(y)) \left| \frac{\partial g_2^{-1}(y)}{\partial y} \right| \mathbf{1}_\wedge(y) \\
&= \frac{1}{\sqrt{2\pi}} \exp(-\frac{y}{2}) \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} \exp(-\frac{y}{2}) \frac{1}{2\sqrt{y}} \qquad (y > 0) \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} \exp(-\frac{y}{2})
\end{aligned} \tag{5}
$$

The 'trick' was to split up the variable transformation in two parts to adjust for the fact that the space of the chi-squared and the Normal are different. We can reverse the same procedure to get from a chi-squared to a normal distribution. We keep the variable names from before. Let $X = \sqrt{Y}$ and therefore $h(x) = \sqrt{x}$. Then $h_1^{-1} : \mathbb{R} \to (-\infty, 0)$ by $h_1^{-1}(x) = -x^2$ and $h_2^{-1} : \mathbb{R} \to [0, \infty)$ by $h_2^{-1}(x) = x^2$. Then

$$
\left| \frac{\partial h_i^{-1}(y)}{\partial y} \right| = |2x|
$$

and

$$
\begin{aligned}
f_X(x) &= f_y(h_1^{-1}(x)) \left| \frac{\partial h_1^{-1}(y)}{\partial y} \right| \mathbf{1}_\wedge(y) + f_y(h_2^{-1}(x)) \left| \frac{\partial h_2^{-1}(y)}{\partial y} \right| \mathbf{1}_\wedge(y) \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{2\sqrt{x^2}} \exp(-\frac{x^2}{2}) |2x| \mathbf{1}_{(-\infty,0)}(x) + \frac{1}{\sqrt{2\pi}} \frac{1}{2\sqrt{x^2}} \exp(-\frac{x^2}{2}) |2x| \mathbf{1}_{[0,\infty)}(x) \\
&= \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})
\end{aligned} \tag{6}
$$

which is defined on the entirety of $\mathbb{R}$.

1 Gamma Distribution

1.1 Standard Gamma Distribution

$$
f(x, \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot x^{(\alpha-1)} \cdot e^{(-\lambda x)} \tag{7}
$$

where $\Gamma(\alpha)$ is the Gamma function. This can be written as

$$
f(x, \alpha, \lambda) = \exp\left[(\alpha - 1)\log(x) - \lambda x + \alpha \log(\lambda) - \log(\Gamma(\alpha))\right] \tag{8}
$$

$$
= \frac{1}{x} \exp\left[\alpha \log(x) - \lambda x + \alpha \log(\lambda) - \log(\Gamma(\alpha))\right] \tag{9}
$$

with $h(x) = \frac{1}{x}, T = (\log x, x), \eta = (\alpha, -\lambda)$ and $A(\alpha, \lambda) = \log(\Gamma(\alpha)) - \alpha \log(\lambda)$.

1.1.1 Laplace Approximation of the Gamma Distribution

To get the LPA of the Gamma function in the standard basis we need its mode and the second derivative of the log-pdf. The mode is already known to be $\hat{\theta} = \frac{\alpha-1}{\lambda}$. For the second derivative of the log-pdf we take the log-pdf, derive it twice and insert the mode for x:

$$\begin{aligned}
\text{log-pdf:}\ & \log\left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\cdot x^{(\alpha-1)}\cdot e^{(-\lambda x)}\right) \\
& = \alpha\cdot\log(\lambda) - \log(\Gamma(\alpha)) + (\alpha-1)\log(x) - \lambda x \\
\text{1st derivative:}\ & \frac{(\alpha-1)}{x} - \lambda \\
\text{mode:}\ & \frac{(\alpha-1)}{x} - \lambda = 0 \Leftrightarrow x = \frac{\alpha-1}{\lambda} \\
\text{2nd derivative:}\ & -\frac{(\alpha-1)}{x^2} \\
\text{insert mode:}\ & -\frac{(\alpha-1)}{(\frac{\alpha-1}{\lambda})^2} = -\frac{\lambda^2}{\alpha-1} \\
\text{invert and times -1:}\ & \sigma^2 = \frac{\alpha-1}{\lambda^2}
\end{aligned}$$

The LPA of the Gamma distribution is therefore approximately distributed according to the pdf of $\mathcal{N}(\frac{\alpha-1}{\lambda}, \frac{\lambda^2}{\alpha-1})$.

1.2 Sqrt-Transform of the Gamma Distribution

1.2.1 Sqrt-Transformation

We transform the Gamma Distribution with the sqrt-transformation, i.e. $Y = \sqrt{X}, g(x) = \sqrt{x}, g_1^{-1}(y) = -y^2, g_2^{-1}(y) = y^2$ and $\left|\frac{\partial g_i^{-1}(y)}{\partial y}\right| = |2x|$. We use the same 'trick' as in Subsection 0.3 to split up the transformation in two parts.

$$f_X(x,\alpha,\lambda) = \frac{1}{x}\exp\left[\alpha\log(x) - \lambda x + \alpha\log(\lambda) - \log(\Gamma(\alpha))\right] \tag{10}$$

$$\begin{aligned}
f_Y(y) &= f_X(g_1^{-1}(y))\left|\frac{\partial g_1^{-1}(y)}{\partial y}\right|\mathbf{1}_\wedge(y) + f_X(g_2^{-1}(y))\left|\frac{\partial g_2^{-1}(y)}{\partial y}\right|\mathbf{1}_\wedge(y) \\
&= \frac{1}{2y}\exp[\alpha\log(y) - \lambda y - A(\alpha,\lambda)]|2y|\mathbf{1}_{(-\infty,0)}(y) + \frac{1}{2y}\exp[\alpha\log(y) - \lambda y - A(\alpha,\lambda)]|2y|\mathbf{1}_{[0,\infty)}(y) \\
&= \exp[\alpha\log(y) - \lambda y - A(\alpha,\lambda)]\mathbf{1}_{(-\infty,+\infty)}(y) \qquad (11) \\
&= \exp[\alpha\log(y) - \lambda y - A(\alpha,\lambda)]
\end{aligned}$$

which is defined on the entirety of $\mathbb{R}$ and is an exponential family with $h(y) = 1, T = (\log y, y), \eta = (\alpha, -\lambda)$ and $A(\alpha,\lambda) = \log(\Gamma(\alpha)) - \alpha\log(\lambda)$.

1.2.2 Laplace Approximation of the sqrt-transformed Gamma Distribution

To get the LPA of the Gamma distribution in the transformed basis we need to calculate its mode and the second derivative of the log-pdf w.r.t x (!). The easiest way to do that is to replace all y by x^2. After inserting $y = x^2$ the pdf does not integrate to 1 anymore and we have to recalculate the integral. This is not a necessary step, since proportionality is sufficient for our purposes.

$$
\begin{aligned}
f_Y(x^2) &= x^{2\alpha} - \exp(-\lambda x^2) \cdot \frac{1}{C} && (12)\\
&= x^{2\alpha} - \exp(-\lambda x^2) \cdot \frac{l^{\alpha+\frac{1}{2}}}{\Gamma\left(\frac{2\alpha+1}{2}\right)} && (13)\\
&= \exp\left[2\alpha\log(x) - \lambda 2x + (\alpha+0.5)\log(\lambda) - \log\left(\Gamma\left(\frac{2\alpha+1}{2}\right)\right)\right] && (14)\\
& && (15)
\end{aligned}
$$

since

$$
C = \int_x x^{2\alpha} - \exp(-\lambda x^2) = l^{-\alpha-\frac{1}{2}} \cdot \Gamma\left(\frac{2\alpha+1}{2}\right) = \frac{\Gamma\left(\frac{2\alpha+1}{2}\right)}{l^{\alpha+\frac{1}{2}}} \tag{16}
$$

To get the mode we take the first derivative and set it to zero.

$$
\begin{aligned}
\text{log-pdf: } & 2\alpha\log(x) - \lambda 2x - C && (17)\\
\text{1st derivative: } & \frac{2\alpha}{x} - 2\lambda x && (18)\\
\text{mode: } & \frac{2\alpha}{x} - 2\lambda x = 0 \Leftrightarrow x = \sqrt{\frac{\alpha}{\lambda}} && (19)\\
\text{2nd derivative: } & -\frac{2\alpha}{x^2} - 2\lambda && (20)\\
\text{insert mode: } & -\frac{2\alpha}{\frac{\alpha}{\lambda}} - 2\lambda = -4\lambda && (21)\\
\text{invert and times -1: } & \frac{1}{4\lambda} && (22)
\end{aligned}
$$

Therefore the LPA now is $\mathcal{N}\left(\sqrt{\frac{\alpha}{\lambda}}, \frac{1}{4\lambda}\right)$.

1.2.3 The bridge for the sqrt-transformation

We already know how to get μ and σ from λ and α. To invert we calculate $\mu = \sqrt{\frac{\alpha}{\lambda}} \Leftrightarrow \alpha = \frac{\mu^2}{\lambda}$ and insert $\lambda = \frac{4}{\sigma^2}$. In summary we have

$$
\begin{aligned}
\mu &= \sqrt{\frac{\alpha}{\lambda}} && (23)\\
\sigma^2 &= \frac{1}{4\lambda} && (24)\\
\lambda &= \frac{4}{\sigma^2} && (25)\\
\alpha &= \frac{(\sigma\mu)^2}{4} && (26)
\end{aligned}
$$

1.3 Log-Transform of the Gamma Distribution

1.3.1 Log-Transformation

We transform the Gamma Distribution with the Log-Transformation, i.e. $Y = \log(X), g(x) = \log(x), g^{-1}(x) = \exp(x)$. The transformed pdf is

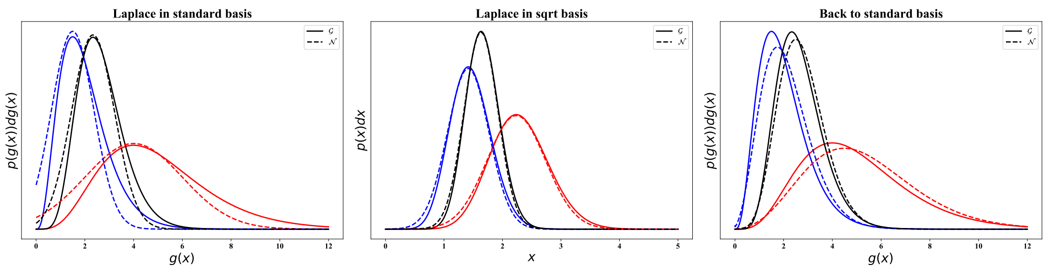

Figure 1: gamma comparison sqrt

$$f_Y(y, \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot y^{(\alpha-1)} \cdot e^{(-\lambda y)} \cdot y \tag{27}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot y^{\alpha} \cdot e^{(-\lambda y)} \tag{28}$$

which can be rewritten as

$$f_Y(y, \alpha, \lambda) = \exp\left[\alpha \log(y) - \lambda \exp(y)\alpha \log(\lambda) - \log(\Gamma(\alpha))\right] \tag{29}$$

with $T = (\log(y), y), \eta = (\alpha, -\lambda)$ and $A(\alpha, \lambda) = \log(\Gamma(\alpha)) - \alpha \log(\lambda)$.
If we insert $y = \exp(x)$ we get

$$f_Y(x, \alpha, \lambda) = \cdot \exp(x)^\alpha \cdot e^{(-\lambda \exp(x))} \cdot \frac{1}{C} \tag{30}$$

$$= \exp\left[\alpha x - \lambda \exp(x) - C\right] \tag{31}$$

where C is the integral but has no nice closed form solution.

1.3.2 Laplace Approximation of the log-transformed Gamma Distribution

To get the LPA of the Gamma distribution in the transformed basis we need to calculate its mode and the second derivative of the log-pdf. To get the mode we take the first derivative and set it to zero.

$$\text{log-pdf: } \log\left(1/C \cdot \exp(x)^\alpha \cdot \exp(-\lambda \exp(x))\right) \tag{32}$$

$$= -C + \alpha x - \lambda \exp(x) \tag{33}$$

$$\text{1st derivative: } \alpha - \lambda \exp(x) \tag{34}$$

$$\text{mode: } \alpha - \lambda \exp(x) = 0 \Leftrightarrow x = \log\left(\frac{\alpha}{\lambda}\right) \tag{35}$$

$$\text{2nd derivative: } -\lambda \exp(x) \tag{36}$$

$$\text{insert mode: } -\lambda \exp(\log\left(\frac{\alpha}{\lambda}\right)) = -\alpha \tag{37}$$

$$\text{invert and times -1: } \sigma^2 = 1/\alpha \tag{38}$$

Therefore the LPA now is $N(\log\left(\frac{\alpha}{\lambda}\right), \alpha)$.

1.3.3 The bridge for the log-transformation

We already know how to get μ and σ from λ and α. To invert we calculate $\mu = \log(\alpha/\lambda) \Leftrightarrow \lambda = \alpha/\exp(\mu)$ and insert $\alpha = \sigma^2$. In summary we have

$$
\begin{aligned}
\mu &= \log(\alpha/\lambda) && (39)\\
\sigma^2 &= \alpha && (40)\\
\lambda &= \alpha/\exp(\mu) && (41)\\
\alpha &= 1/\sigma^2 && (42)
\end{aligned}
$$

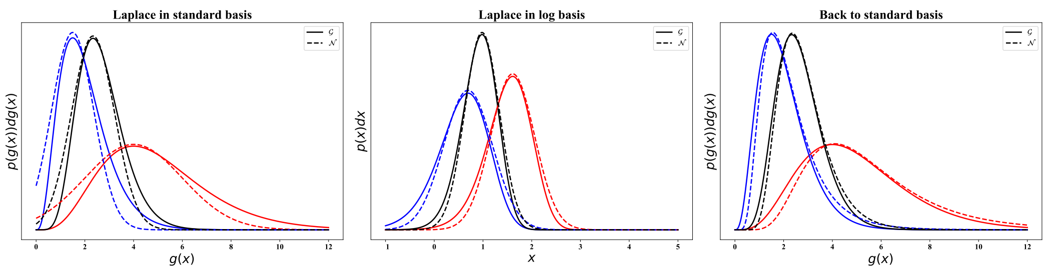

Figure 2: gamma comparison log