Deriving a Gaussian Prior from a Generalized Gamma Hyperprior

Yash Dave

July 5, 2024

Abstract

The broad idea is that we can identify sets of the parameters r and η such that our hierarchical Bayesian model with generalized gamma hyperpriors becomes a Gaussian prior. Here, we derive η as a function of r wherein, in the limit as r and η tends to infinity, we obtain a Gaussian prior with a target variance of our choice.

1 Introduction

Note that we model $x \sim N(0, \theta)$ where $\theta \sim GenGamma(r, \beta, \vartheta)$. That is, x is conditionally Gaussian with mean 0 and variance drawn from a generalized gamma distribution parameterized by two shape parameters r and η, and one scale parameter $\vartheta = 1$.

For the prior model to collapse into a Gaussian prior with a target variance $\bar{\theta}$, we need $\mathbf{E}[\theta] = \bar{\theta}$ and $\mathbf{Var}[\theta] \to 0$. The broad approach is as follows:

1. Substitute $\beta = \frac{\eta+1.5}{r}$ for clarity
2. Let β be a function of r and implicitly differentiate to obtain an ODE in terms of $\beta'(r) = \frac{d\beta}{dr}$
3. Apply asymptotic limits to obtain an approximate solution for $\beta(r)$

By interchanging the order of integrals as in the Generalized Gamma Notes, we obtain the following formulas:

$$\mathbf{E}[\theta] = \Gamma(\frac{\eta + 2.5}{r})/\Gamma(\frac{\eta + 1.5}{r}) = \frac{\Gamma(\beta + r^{-1})}{\Gamma(\beta)} \tag{1}$$

$$\mathbf{Var}[\theta] = \Gamma(\frac{\eta + 3.5}{r})/\Gamma(\frac{\eta + 1.5}{r}) = \frac{\Gamma(\beta + 2r^{-1})}{\Gamma(\beta)} \tag{2}$$

Focusing on (1), and explicitly writing out $\beta(r)$, we can rearrange to obtain

$$\Gamma(\beta(r) + r^{-1}) = \bar{\theta} \cdot \Gamma(\beta(r)) \tag{3}$$

Letting Γ' denote the derivative of the Gamma function and applying chain rule, we obtain

$$\Gamma'(\beta(r) + r^{-1}) \cdot (\beta'(r) - r^{-2}) = \bar{\theta} \cdot \Gamma'(\beta(r)) \cdot \beta'(r) \tag{4}$$

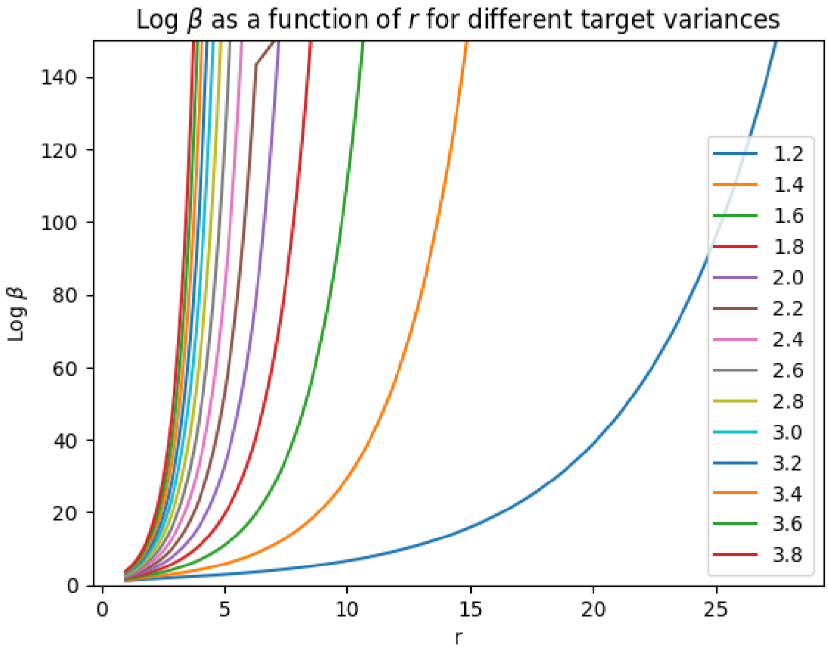
$$\Gamma'(\beta(r) + r^{-1}) \cdot \beta'(r) - \bar{\theta} \cdot \Gamma'(\beta(r)) \cdot \beta'(r) = r^{-2}\Gamma'(\beta(r) + r^{-1}) \tag{5}$$

$$\beta'(r) = r^{-2} \frac{\Gamma'(\beta(r) + r^{-1})}{\Gamma'(\beta(r) + r^{-1}) - \bar{\theta} \cdot \Gamma'(\beta(r))} \tag{6}$$

For clarity, we drop the r when referring to $\beta(r)$ and rearrange to obtain

$$\beta' = r^{-2}\Big(1 - \bar{\theta}\frac{\Gamma'(\beta)}{\Gamma'(\beta + r^{-1})}\Big)^{-1} \tag{7}$$

Visually, this is the ODE whose level sets give rise to the curves in the plot below:



In order to proceed further, we examine the derivative of the gamma function. Generally, observe that

$$\frac{d}{dx}\log(f(x)) = \frac{f'(x)}{f(x)} \tag{8}$$

In the case of the Gamma function, this helps us define the Digamma function below

$$\Psi_0(z) = \frac{\Gamma'(z)}{\Gamma(z)} \tag{9}$$

Equivalently,

$$\Psi_0(z)\Gamma(z) = \Gamma'(z) \tag{10}$$

In terms of Ψ_0, equation (12) becomes

$$\beta' = r^{-2}\Big(1 - \bar{\theta}\frac{\Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\frac{\Gamma(\beta)}{\Gamma(\beta + r^{-1})}\Big)^{-1} = r^{-2}\Big(1 - \bar{\theta}\frac{\Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\frac{1}{\bar{\theta}}\Big)^{-1} \tag{11}$$

Altogether, this simplifies to

$$\beta' = r^{-2}\Big(1 - \frac{\Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\Big)^{-1} \tag{12}$$

Intuitively, to understand the behaviour of β as r goes to infinity, we can see it's likely the case that β is an increasing function that grows superlinearly. This is evident in the plot earlier, but equally so by recognizing that Gamma is an increasing function.

In order to get a better bound of sorts, we consider the 1st-degree Taylor approximation of $\Psi_0(\beta + r^{-1}$ about β.

$$\Big(1 - \frac{\Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\Big) = \Big(\frac{\Psi_0(\beta + r^{-1}) - \Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\Big) \tag{13}$$

$$\approx \Big(\frac{\Psi_0(\beta) + \Psi_0'(\beta)((\beta + r^{-1}) - \beta) - \Psi_0(\beta)}{\Psi_0(\beta) + \Psi_0'(\beta)(\beta - (\beta + r^{-1}))}\Big) \tag{14}$$

$$= \Big(\frac{\Psi_0'(\beta) r^{-1}}{\Psi_0(\beta) + \Psi_0'(\beta)(r^{-1})}\Big) \tag{15}$$

Using the approximation that $\frac{1}{1+x} \approx 1 - x$ for small x, here for $x = \frac{\Psi_0'(\beta) r^{-1}}{\Psi_0(\beta)}$ we conclude

$$\Big(1 - \frac{\Psi_0(\beta)}{\Psi_0(\beta + r^{-1})}\Big) \approx \frac{\Psi_0'(\beta) r^{-1}}{\Psi_0(\beta)}\Big(1 - \frac{\Psi_0'(\beta) r^{-1}}{\Psi_0(\beta)}\Big) \tag{16}$$

$$\approx r^{-1}\frac{\Psi_0'(\beta)}{\Psi_0(\beta)} + \text{higher order terms} \tag{17}$$

$$\approx r^{-1}\frac{\Psi_1(\beta)}{\Psi_0(\beta)} \tag{18}$$

where Ψ_1 is the Trigamma function defined to be Ψ_1.
Plugging this expression back into (12), we obtain

$$\beta' \approx r^{-2}\Big(r^{-1}\frac{\Psi_1(\beta)}{\Psi_0(\beta)}\Big)^{-1} = r^{-1}\frac{\Psi_0(\beta)}{\Psi_1(\beta)} \tag{19}$$

Using Wolfram to compute the Puiseux series of the ratio $\Psi_1(x)/\Psi_0(x)$ at $x = \infty$, we obtain

$$\beta' \approx r^{-2}\Big(r^{-1}\frac{\Psi_1(\beta)}{\Psi_0(\beta)}\Big)^{-1} \tag{20}$$

$$\approx r^{-1}(\beta \log(\beta) + O(\beta^{-2})) \tag{21}$$

$$\approx \frac{\beta \log(\beta)}{r} \tag{22}$$

The ODE is now in a much cleaner form and we can attempt to solve it. Recall the definition of the logarithmic derivative from (8). In this context, let

$$y = \frac{\beta'}{\beta} = \frac{d}{dr}\log(\beta) \tag{23}$$

This substitution then results in the following ODE

$$y' = \frac{y}{r} \tag{24}$$

Once again using the same trick, note that

$$\frac{d}{dr}\log(y) = \frac{1}{r} \tag{25}$$

Since the equation is separable, we can now solve for $y(r)$ by integrating both sides with respect to r to give

$$\log(y) = \int \frac{1}{r} dr = \log(r) + c_0 \tag{26}$$

Note that r is a positive value as we are looking at the limit as r tends to infinity, so there is no need for an absolute value.

Exponentiating both sides,

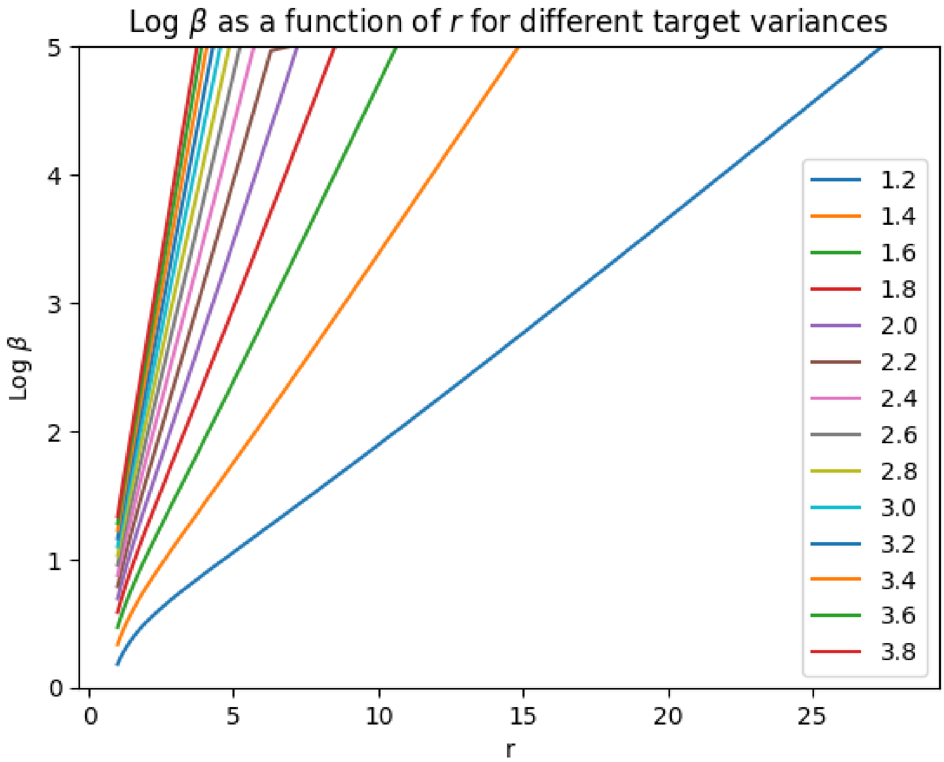
$$\log \beta = y = \int \frac{1}{r} dr = e^{\log(r)+c_0} = c_1 r \tag{27}$$

Finally,

$$\beta(r) = e^{c_1 r} = c_2 e^r \tag{28}$$

We can set our initial values by picking a value r_0 and computing $\beta_0 = \beta(r_0)$. The resulting manifold will give rise to a Gaussian distribution with variance given by equation (1) with inputs (r_0, β_0).

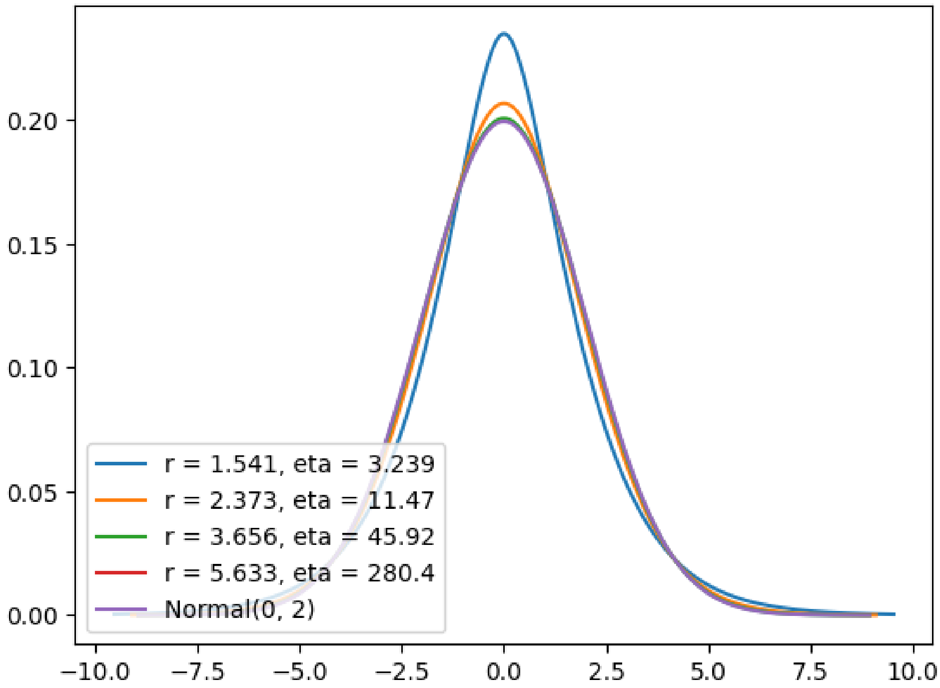
This tells us that β grows exponentially in r. In turn, this implies that $\log(\beta)$ against r produces a linear plot, as confirmed below:



In terms of η, we find

$$\eta(r) = rce^{r} + O(r) \tag{29}$$

For a visual sense of what it means to traverse this manifold, consider the plot below:



Note how η grows superlinearly as compared to r in order to ensure that the target variance is achieved.