higher order rendering equations

zican wang

August 2024

1 Derivation for gradient and hessian of n-dimensional isotropic mean zero Gaussian $G(\mathbf{x})$:

Higher dimensional Gaussians are separable:

$$
\begin{aligned}
G_{\mathrm{ND}}(\mathbf{x};\sigma) &= \frac{1}{(\sqrt{2\pi}\sigma)^N} e^{-\frac{|\mathbf{x}|^2}{2\sigma^2}} && (1)\\
&= \underbrace{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{{x_1}^2}{2\sigma^2}} \times \cdots \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{{x_i}^2}{2\sigma^2}} \times \cdots}_{1\ldots N} && (2)\\
&= G_{\mathrm{1D}}(x_1;\sigma) \times \cdots \times G_{\mathrm{1D}}(x_N;\sigma) && (3)\\
&= \prod_{i=1}^{N} G_{\mathrm{1D}}(x_i;\sigma) && (4)
\end{aligned}
$$

For gradient in one dimension, $G_{\mathrm{1D}}(x_i)$ is short hand for $G_{\mathrm{1D}}(x_i;\sigma)$:

$$
\begin{aligned}
\frac{dG_{\mathrm{1D}}(x_i)}{dx_i} &= \frac{d}{dx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{{x_i}^2}{2\sigma^2}} && (5)\\
&= -\frac{x_i}{\sigma^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{{x_i}^2}{2\sigma^2}} && (6)\\
&= -\frac{x_i}{\sigma^2} G_{\mathrm{1D}}(x_i) && (7)
\end{aligned}
$$

For higher dimensions:

$$
\begin{aligned}
\frac{\partial G_{\mathrm{ND}}}{\partial x_i} &= \frac{dG_{\mathrm{1D}}(x_i)}{dx_i} \prod_{\substack{j=1\\ j\neq i}}^{N} G_{\mathrm{1D}}(x_j) && (8)\\
&= -\frac{x_i}{\sigma^2} G_{\mathrm{1D}}(x_i) \prod_{\substack{j=1\\ j\neq i}}^{N} G_{\mathrm{1D}}(x_j) && (9)\\
&= -\frac{x_i}{\sigma^2} G_{\mathrm{ND}}(\mathbf{x}) && (10)
\end{aligned}
$$

For diagonal of hessian:

$$\frac{\partial^2 G_{\mathrm{ND}}}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(-\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\prod_{\substack{j=1\\j\neq i}}^{N}G_{1\mathrm{D}}(x_j)\right) \quad \text{(from Eq.9)} \tag{11}$$

$$= \frac{\partial}{\partial x_i}\left(-\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\right)\prod_{\substack{j=1\\j\neq i}}^{N}G_{1\mathrm{D}}(x_j) \tag{12}$$

$$= \left(-\frac{1}{\sigma^2}+\frac{{x_i}^2}{\sigma^4}\right)G_{1\mathrm{D}}(x_i)\prod_{\substack{j=1\\j\neq i}}^{N}G_{1\mathrm{D}}(x_j) \tag{13}$$

$$= \left(-\frac{1}{\sigma^2}+\frac{{x_i}^2}{\sigma^4}\right)G_{\mathrm{ND}}(\mathbf{x}) \tag{14}$$

For non-diagonal of hessian:

$$\frac{\partial^2 G_{\mathrm{ND}}}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(-\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\prod_{\substack{k=1\\k\neq i}}^{N}G_{1\mathrm{D}}(x_k)\right) \quad \text{(from Eq.9)} \tag{15}$$

$$= \left(-\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\right)\left(-\frac{x_j}{\sigma^2}G_{1\mathrm{D}}(x_j)\right)\prod_{\substack{k=1\\k\neq i,j}}^{N}G_{1\mathrm{D}}(x_k) \tag{16}$$

$$= \frac{x_i x_j}{\sigma^4}G_{1\mathrm{D}}(x_i)G_{1\mathrm{D}}(x_j)\prod_{\substack{k=1\\k\neq i,j}}^{N}G_{1\mathrm{D}}(x_k) \tag{17}$$

$$= \frac{x_i x_j}{\sigma^4}G_{\mathrm{ND}}(\mathbf{x}) \tag{18}$$

2 Derivation of IS sampler for gradient of mean zero Gaussian $G(\mathbf{x})$:

Inverse transform sampling is used to get the samples from the given distribution. For this to work, we need to compute a valid CDF and PDF of the gradient of the Gaussian. If we want the gradient to be a valid distribution, we need to positivise it and multiply by some factor α to make its CDF monotonically increasing and integral to 1.

2.1 For 1-dimensional Gaussian:

For 1D gradient, we have $-\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)$, the integral of the gradient with constant 0 would return the Gaussian function. We can see that the 1D gradient is negative when $x_i > 0$, which means that the peak of the Gaussian function(its integral) is at $x_i = 0$ (Figure 1, blue and orange). To positivise it, all negative values are turned into positive values with same size.

$$\frac{\partial G}{\partial x_i} = \begin{cases} -\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i) \geq 0 & : x_i \leq 0 \\ -\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i) < 0 & : x_i > 0 \end{cases} \Rightarrow \text{positivised} \begin{cases} -\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i) \geq 0 & : x_i \leq 0 \\ \frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i) > 0 & : x_i > 0 \end{cases} \tag{19}$$

After positivisation, we will have a symmetric curve about the y axis(Figure 1, green). Thus, the two parts should contribute equally to the CDF, meaning that at $x_i = 0$, the CDF value should be 0.5. For $x_i > 0$, the non-positivised gradient is negative, subtracting from 0.5. To adapt to the positivisation, the subtracted amount $0.5 - G_{1\mathrm{D}}(x_i)$ will be added to 0.5 instead.

$$\text{Integral} \begin{cases} \int_{-\infty}^{\infty} -\frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\,dx_i & : x_i \leq 0 \\ \int_{-\infty}^{\infty} \frac{x_i}{\sigma^2}G_{1\mathrm{D}}(x_i)\,dx_i & : x_i > 0 \end{cases} \Rightarrow \text{CDF} \begin{cases} 0 + \alpha G_{1\mathrm{D}}(x_i) \leq 0.5 & : x_i \leq 0 \\ 0.5 + 0.5 - \alpha G_{1\mathrm{D}}(x_i) > 0.5 & : x_i > 0 \end{cases} \tag{20}$$

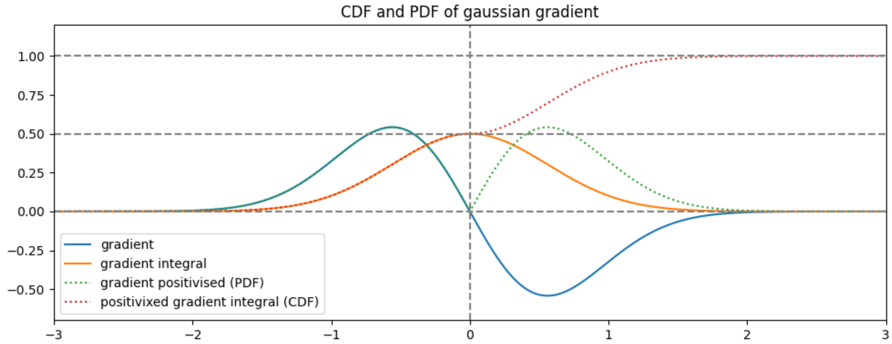

Figure 1: The PDF and CDF of gradient of 1D Gaussian

Finally, we can solve for α by making sure the CDF value is 0.5 at $x_i = 0$.

$$\alpha G_{1\text{D}}(0) = 1 - \alpha G_{1\text{D}}(0) = 0.5 \tag{21}$$

$$\alpha \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{0^2}{2\sigma^2}} = 0.5 \tag{22}$$

$$\alpha' e^{-\frac{0^2}{2\sigma^2}} = 0.5 \tag{23}$$

$$\alpha' = 0.5 \tag{24}$$

Thus we have:

$$\frac{dG_{1\text{D}}}{dx_i}^{\text{CDF}} \begin{cases} 0.5e^{-\frac{x_i{}^2}{2\sigma^2}} & : x_i \leq 0 \\ 1 - 0.5e^{-\frac{x_i{}^2}{2\sigma^2}} & : x_i > 0 \end{cases}, \quad \frac{dG_{1\text{D}}}{dx_i}^{\text{PDF}} = \alpha \frac{|x_i|}{\sigma^2} G_{1\text{D}}(x_i) = \frac{|x_i|}{2\sigma^2} e^{-\frac{x_i{}^2}{2\sigma^2}} \tag{25}$$

And for inverse of CDF:

$$\frac{dG_{1\text{D}}}{dx_i}^{\text{ICDF}} \begin{cases} -\sqrt{2\sigma^2 \log(2\xi)} & : x_i \leq 0, \xi \leq 0.5 \\ \sqrt{2\sigma^2 \log(2(1-\xi))} & : x_i > 0, \xi > 0.5 \end{cases} \tag{26}$$

2.2 For n-dimensional isotropic Gaussian:

For the partial derivative $\frac{\partial G}{\partial x_i}$, the gradients are shown in Equation 9. As there are multiple dimensions in the space of $\mathbf{x}$, each dimension is sampled one by one from the marginalization of the selected dimension, given the already sampled dimension(Figure 2). Since the Gaussian and its derivatives are separable, the marginalisation of any dimension would be the same as a slice of that dimension(given the already sampled dimensions), once they are normalised into a probability distribution to sample from.

Marginalizing over x_i:

$$\int_{x_{j\neq i}} \frac{\partial G}{\partial x_i}\, dx_{j\neq i} = -\frac{x_i}{\sigma^2} G_{1\text{D}}(x_i) \int_{x_{j\neq i}} \prod_{\substack{j=1 \\ j\neq i}}^{N} G_{1\text{D}}(x_j)\, dx_{j\neq i} \tag{27}$$

$$= -\frac{x_i}{\sigma^2} G_{1\text{D}}(x_i) \int G_{1\text{D}}(x_1) \cdots \int G_{1\text{D}}(x_{i-1}) \int G_{1\text{D}}(x_{i+1}) \cdots \int G_{1\text{D}}(x_N)\, dx_N \cdots dx_1 \tag{28}$$

$$= -\frac{x_i}{\sigma^2} G_{1\text{D}}(x_i) \tag{29}$$

$$\tag{30}$$

This is the same as the 1-dimensional gradient, which means that the distribution from positivisation and re-scaling is also the same. The PDF of the joint distribution would be:

$$\frac{\partial G_{\text{ND}}}{\partial x_i}^{\text{PDF}} = \frac{\partial G_{1\text{D}}}{\partial x_i}^{\text{PDF}} \prod_{\substack{j=1 \\ j\neq i}}^{N} G_{1\text{D}}(x_j) \tag{31}$$

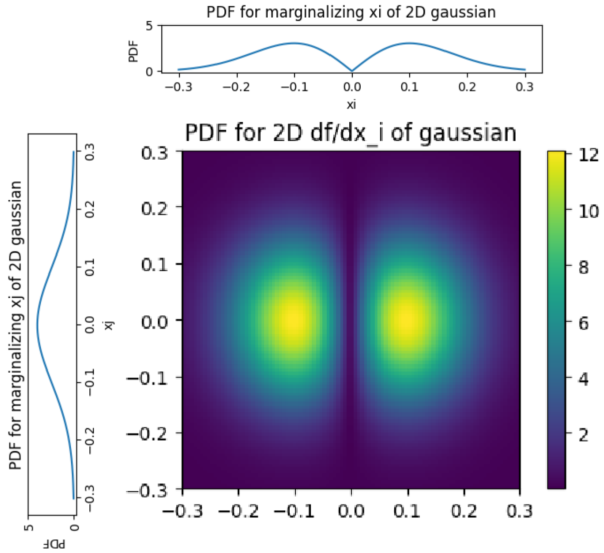

Figure 2: The PDF and marginalization of gradient of 2D Gaussian, if we first sample from x_i, we use the marginalization of x_i for the PDF to get a x_i value, then sample x_j in the selected column given x_i

Since each dimension is a valid distribution so their integral would be 1, and the distribution is separable, marginalizing over any x_j would result in the 1-dimensional version of the Gaussian distribution:

$$\int_{x_{k\neq j}} \frac{\partial G_{\text{ND}}}{\partial x_i}\, dx_{k\neq j} = G_{\text{1D}}(x_j) \int_{x_{k\neq j}} \frac{\partial G_{\text{1D}}}{\partial x_i}^{\text{PDF}} \prod_{\substack{k=1 \\ k\neq i,j}}^{N} G_{\text{1D}}(x_k)\, dx_{k\neq i} \tag{32}$$

$$= G_{\text{1D}}(x_j) \tag{33}$$

In conclusion, for n-dimensional Gaussian gradient, we can sample each dimension independently from corresponding distributions, $\frac{\partial G_{\text{1D}}}{\partial x_i}^{\text{DIST.}}$ for the dimension of the partial derivative x_i and Gaussian for the rest. From equation 27, we can also see that sampling from 1-dimensional Gaussians for all $n-1$ dimensions independently is equivalent to sampling from a $(n-1)$ dimensional Gaussian. The corresponding PDF function of the joint distribution would be the product of the PDF for 1 dimensional Gaussian gradient and PDF for $n-1$ dimensional Gaussian.

$$\frac{\partial G_{\text{ND}}}{\partial x_i}^{\text{PDF}} = \frac{\partial G_{\text{1D}}}{\partial x_i}^{\text{PDF}} \prod_{\substack{j=1 \\ j\neq i}}^{N} G_{\text{1D}}(x_j) \tag{34}$$

$$= \frac{\partial G_{\text{1D}}}{\partial x_i}^{\text{PDF}} G_{(N-1)\text{D}}(\mathbf{x}_{\setminus i}) \tag{35}$$

3 Derivation of IS sampler for Hessian of mean zero Gaussian:

The hessian of a n-dimensional Gaussian have two different functions for the diagonal and non-diagonal.

3.1 For diagonal of Hessian/second order derivative:

For the diagonal of the hessian(Eq.13), we can follow similar logic and sample each dimension independently. Thus we can first derive the valid distribution of the second order derivative of the 1-dimensional Gaussian by positivisation and scaling(Figure 3), and it will apply to higher dimensions:

Solving for roots for the 1-dimensional Gaussian's second order derivative $\left(-\frac{1}{\sigma^2} + \frac{{x_i}^2}{\sigma^4}\right) G_{\text{1D}}(x_i) = 0$, we get the values $x = \pm\sigma$. Thus, the function value between $-\sigma < x \leq \sigma$ should be positivised. Since the second order derivative should integrate to the gradient of the Gaussian, we know that it

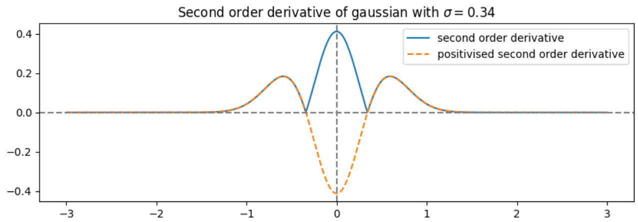

Figure 3: Positivised 2nd order derivative of Gaussian

reaches zero as x_i reaches infinity. In addition to the fact that the the second order derivative is symmetric about the y axis, we can conclude that the integral of the interval $-\sigma < x \leq \sigma$ should be twice the size of the integral of $x \leq -\sigma / x > \sigma$. Thus, after positivisation and scaling, the CDF should be 0.25 at $x_i = -\sigma$ and 0.75 at $x_i = \sigma$ (Figure 4). Solving for these equalities, we can get:

$$\beta \frac{\partial G_{1\mathrm{D}}}{\partial x_i}(x_i) = 0.5 - \beta \frac{\partial G_{1\mathrm{D}}}{\partial x_i}(x_i) = \frac{1}{4} \tag{36}$$

$$\beta \left(-\frac{x_i}{\sigma^2} G_{1\mathrm{D}}(x_i) \right) = \frac{1}{4}, \quad \beta G_{1\mathrm{D}}(x_i) = \beta' \left(e^{-\frac{x_i^2}{2\sigma^2}} \right) \tag{37}$$

$$-\frac{x_i}{\sigma^2} \beta' \left(e^{-\frac{x_i^2}{2\sigma^2}} \right) = \frac{1}{4}, \quad x_i = -\sigma \tag{38}$$

$$\beta' = \frac{\sigma}{4} e^{\frac{1}{2}} \tag{39}$$

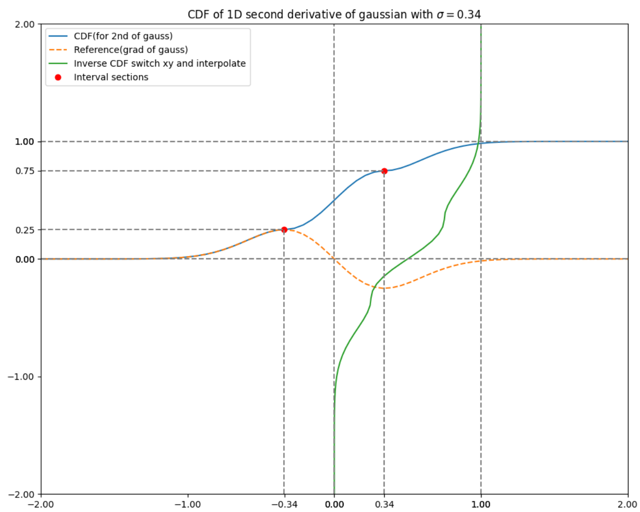

Figure 4: CDF and ICDF of Gaussian second order derivative

So, for the positivised rescaled second order derivative as a PDF of the distribution:

$$\frac{d^2 G_{1\mathrm{D}}}{dx_i^2}^{\mathrm{PDF}} = \begin{cases} \beta \frac{d^2 G_{1\mathrm{D}}}{dx_i^2} & : x_i \leq -\sigma, x_i \geq \sigma \\ -\beta \frac{d^2 G_{1\mathrm{D}}}{dx_i^2} & : -\sigma < x_i \leq \sigma \end{cases} \tag{40}$$

$$= \begin{cases} \beta \left(-\frac{1}{\sigma^2} + \frac{x_i{}^2}{\sigma^4} \right) G_{1\mathrm{D}}(x_i) & : x_i \leq -\sigma, x_i \geq \sigma \\ -\beta \left(-\frac{1}{\sigma^2} + \frac{x_i{}^2}{\sigma^4} \right) G_{1\mathrm{D}}(x_i) & : -\sigma < x_i \leq \sigma \end{cases} \tag{41}$$

$$= \begin{cases} \left(-\frac{1}{4\sigma} + \frac{x_i{}^2}{4\sigma^3} \right) e^{\frac{1}{2} - \frac{x_i^2}{2\sigma^2}} & : x_i \leq -\sigma, x_i \geq \sigma \\ -\left(-\frac{1}{4\sigma} + \frac{x_i{}^2}{4\sigma^3} \right) e^{\frac{1}{2} - \frac{x_i^2}{2\sigma^2}} & : -\sigma < x_i \leq \sigma \end{cases} \tag{42}$$

We can also get the CDF function for the other intervals by flipping and translating the scaled gradient of Gaussian (Figure 4):

$$\frac{d^2G_{\text{1D}}}{dx_i^2}^{\text{CDF}} = \begin{cases} -\frac{x_i}{4\sigma}e^{\frac{1}{2}-\frac{x_i^2}{2\sigma^2}} & : x_i \leq -\sigma \\ 0.5 + \frac{x_i}{4\sigma}e^{\frac{1}{2}-\frac{x_i^2}{2\sigma^2}} & : -\sigma < x_i \leq \sigma \\ 1 - \frac{x_i}{4\sigma}e^{\frac{1}{2}-\frac{x_i^2}{2\sigma^2}} & : x_i > \sigma \end{cases} \tag{43}$$

However, the is no elementary operation that can represent the analytical form of the inverse of this CDF. Thus the corresponding x_i and its CDF values are precomputed, and for a given ξ as the CDF value, it is linearly interpolated from the precomputed CDF values and the corresponding interpolated x is returned.

3.2 For non-diagonal of Hessian:

As shown in equation 17, the non-diagonal of the hessian for element i, j is equivalent to the product of the gradient in the i and j direction.

$$\frac{\partial^2 G_{\text{ND}}}{\partial x_i \partial x_j} = \frac{dG_{\text{1D}}}{dx_i}\frac{dG_{\text{1D}}}{dx_j}G_{(\text{N}-2)\text{D}}(\mathbf{x}_{\setminus i,j}) \tag{44}$$

As the two gradient functions are independent to each other, to make $\frac{dG_{\text{1D}}}{dx_i}\frac{dG_{\text{1D}}}{dx_j}$ a valid joint distribution, the positivised and re-scaled hessian should have both dimensions being a valid distribution, which then is $\frac{dG_{\text{1D}}}{dx_i}^{\text{PDF}}\frac{dG_{\text{1D}}}{dx_j}^{\text{PDF}}$ (Figure 5). Thus, no extra CDF or ICDF functions need to be derived because the joint distribution is a product of Gaussian for each dimension and two positivised gradients, which is derived in section 2.1.

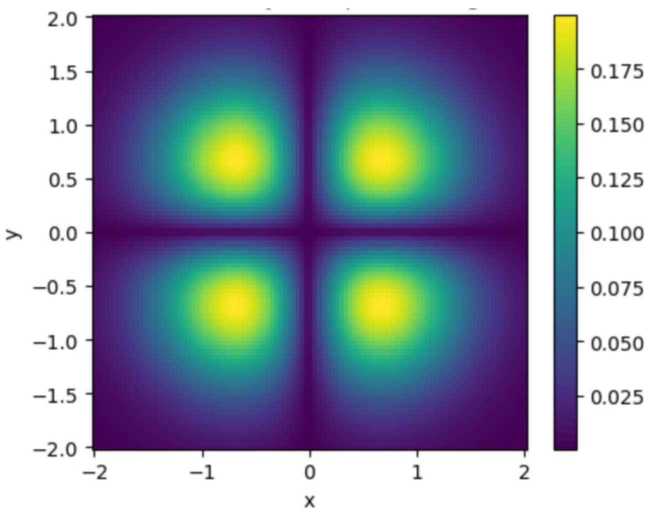

Figure 5: The joint distribution of the non-diagonal hessian