# Score Matching on the Probability Simplex

The last equation required for score matching on the probability simplex is the gradient of the log logit Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi v}} \frac{1}{x(1-x)} \exp\left(-\frac{(\sigma^{-1}(x)-\mu)^2}{2v}\right)$$

where $\sigma^{-1} = \log\left(\frac{x}{1-x}\right)$

We are interested in:
$\nabla_x \log p(x)$
or for the time being:
$\frac{\partial}{\partial x} \log p(x)$

After working in 1D, we will then show the general case.

First we deal with the log prob:

$$\begin{aligned}
\frac{\partial}{\partial x} \log p(x) &= \log\left[\frac{1}{\sqrt{2\pi v}} \frac{1}{x(1-x)} \exp\left(-\frac{(\sigma^{-1}(x)-\mu)^2}{2v}\right)\right] \\
&= C + \log\left[\frac{1}{x(1-x)}\right] - \frac{(\sigma^{-1}(x)-\mu)^2}{2v}
\end{aligned}$$

where $C = \log\left[\frac{1}{\sqrt{2\pi v}}\right]$

Next, we can then differentiate each of the components seperately

The first can be solved as

$$\begin{aligned}
\frac{\partial}{\partial x} \log\left[\frac{1}{x(1-x)}\right] &= \frac{\partial}{\partial x} \log\left[\frac{1}{x}\right] + \frac{\partial}{\partial x} \log\left[\frac{1}{1-x}\right] \\
&= -\frac{\partial}{\partial x} \log[x] - \frac{\partial}{\partial x} \log[1-x] \\
&= -\frac{1}{x} + \frac{1}{1-x}
\end{aligned}$$

and the second can be solved as

$$\begin{aligned}
\frac{\partial}{\partial x} \frac{(\sigma^{-1}(x)-\mu)^2}{2v} &= -\frac{\partial}{\partial x} \frac{(\log\left[\frac{x}{1-x}\right]-\mu)^2}{2v} \\
&= -\frac{\log\left[\frac{x}{1-x}\right]-\mu}{v} \frac{\partial}{\partial x}\left(\log\left[\frac{x}{1-x}\right]-\mu\right) \\
&= -\frac{\log\left[\frac{x}{1-x}\right]-\mu}{vx(1-x)}
\end{aligned}$$

Putting it all together, we get:

$$
\begin{aligned}
\frac{\partial}{\partial x}\log\, p(x) &= \frac{1}{1-x} - \frac{1}{x} + \frac{\mu - \log\, \left[\frac{x}{1-x}\right]}{vx(1-x)} \\
&= \frac{2vx + \mu - v - \log\left[\frac{x}{1-x}\right]}{vx(1-x)} \\
&= \frac{\sigma^{-1}(x) - 2vx - \mu + v}{vx(x-1)}
\end{aligned}
$$

## General Case

The logit-Guassian distribution can be written as:

$$
p(x) = \frac{1}{Z}\frac{1}{\prod_{i=1}^{d} x_i}\exp\, \left( - \frac{\|\log\, \left[\frac{\bar{x}_d}{x_d}\right] - \mu\|_2^2}{2v} \right)
$$

where $x \in \mathcal{S}^d$ and $\bar{x}_d = [x_1, \ldots, x_{d-1}]$.
We assume that the Gaussian has covariance $\Sigma = \sqrt{v} I$

To bring $x$ from the simplex back to $\mathbb{R}^{d-1}$ we can use:

$$
y_i = \log\, \left[\frac{x_i}{x_d}\right], i \in \{1, \ldots, d-1\}
$$

The inverse transformation of this is:

$$
x_i = \frac{e^{y_i}}{1 + \sum_{k=1}^{d-1} e^{y_k}}, i \in \{1, \ldots, d-1\}
$$

$$
x_d = \frac{1}{1 + \sum_{k=1}^{d-1} e^{y_k}} = 1 - \sum_{i=1}^{d-1} x_i
$$

---

Overall, we want to calculate:
$\nabla_x \log\, p(x)$

Following the same process as the 1D case:

$$
\log\, p(x) = -\log\, [Z] - \log\, \left[\prod_{i-1}^{d} x_i\right] - \frac{1}{2v}\|\log\, \left[\frac{\bar{x}_d}{x_d}\right] - \mu\|_2^2
$$

We deal with the gradients, starting with the second term (the first one has no gradient). The fact that

$\log\left[\prod_{i-1}^{d} x_i\right] = \sum_{i=1}^{d} \log\left[x_i\right]$
will be used in the following:

$$g := -\nabla_x \log\left[\prod_{i=1}^{d} x_i\right]$$
$$g_i = -\frac{1}{x_i}$$

—

Next we look at the rightmost term in the equation

$$\nabla_x\left[-\frac{1}{2v}\|\log\left[\frac{\bar{x}_d}{x_d}\right]-\mu\|_2^2\right] := -\frac{1}{2v}\nabla_x\alpha := h$$
$$\nabla_x\alpha = f$$

Now for $i \in \{1, \ldots, d-1\}$ the following holds:

$$\begin{aligned} f_i &= \frac{\partial}{\partial x_i}\left(\log[\frac{x_i}{x_d}]-\mu\right)^2 \\ &= 2\left(\log[\frac{x_i}{x_d}]-\mu\right)\frac{1}{x_i} \\ h_i &= -\frac{\left(\log[\frac{x_i}{x_d}]-\mu\right)}{vx_i} \end{aligned}$$

The last component is more complex to calculate.

$$\begin{aligned} \frac{\partial}{\partial x_d}\alpha &= \frac{\partial}{\partial x_d}\sum_{i=1}^{d-1}\left(\log\left[\frac{x_i}{x_d}\right]-\mu\right)^2 \\ &= 2\sum_{i=1}^{d-1}\left(\log\left[\frac{x_i}{x_d}\right]-\mu\right)\frac{\partial}{\partial x_d}\left(\log\left[\frac{x_i}{x_d}\right]\right) \\ &= 2\sum_{i=1}^{d-1}\left(\log\left[\frac{x_i}{x_d}\right]-\mu\right)\left(-\frac{1}{x_c}\right) \\ h_d &= \frac{d-1}{vx_d}\left(\sum_{i=1}^{d-1}\log\left[\frac{x_i}{x_d}\right]-\mu\right) \end{aligned}$$

In summary, we can then write:

$$
\begin{aligned}
\gamma &:= \nabla_x \log\, p(x) \\
\gamma_i &= -\frac{\log\,[\frac{x_i}{x_d}] - \mu + v}{v x_i}, i \in \{1, \dots, d-1\} \\
\gamma_d &= -\frac{v + (d-1)(\sum_{i=1}^{d-1} \log\,[\frac{x_i}{x_d}] - \mu)}{v x_d}
\end{aligned}
$$