Three forms of the Multivariate Normal

$$\mathcal{N}(\underline{x}; \underline{\mu}, \underline{\underline{\Sigma}}) = \frac{1}{\sqrt{\det(\underline{\underline{\Sigma}})(2\pi)^K}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^T \underline{\underline{\Sigma}}^{-1}(\underline{x}-\underline{\mu})\right)$$

$\underline{x} \in \mathbb{R}^K$

$\underline{\mu} \in \mathbb{R}^K$

$\underline{\underline{\Sigma}} \in \mathbb{R}^{K \times K}$ SPD

reduce footprint?

$$\#\text{params} = K + \cancel{K^2}\ \frac{K \cdot (K+1)}{2} = \mathcal{O}(K^2)$$

$K >> 1000$

① Full Covariance

② Diagonal Covariance

③ Spherical / Isotropical Covariance

---

① Full Covariance

$$\underline{\underline{\Sigma}} = \underline{\underline{\Sigma}}$$

→ allows for defining interactions of the dimensions ("correlation")

in TFP: MultivariateNormalTriL

① Perform a Cholesky decomposition (also check for SPD)

$$\underline{\underline{\Sigma}} = \underline{\underline{L}}\,\underline{\underline{L}}^T$$

② Create RV with $\underline{\mu}$ & $\underline{\underline{L}}$

---

② Diagonal Covariance

$$\underline{\underline{\Sigma}} = \text{diag}(\underline{\sigma}^2)$$

vector of variances

$$\underline{\underline{\Sigma}} = \begin{bmatrix} \sigma_0^2 & & & 0 \\ & \sigma_1^2 & & \\ & & \sigma_2^2 & \\ 0 & & & \ddots \end{bmatrix} \in \mathbb{R}^{K \times K} \qquad \underline{\sigma}^2 = \begin{bmatrix} \sigma_0^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_{K-1}^2 \end{bmatrix} \in \mathbb{R}^K$$

$$\det(\underline{\underline{\Sigma}}) = \prod_{k=0}^{K-1} \lambda_k(\underline{\underline{\Sigma}})$$

"product of eigenvalues"

$$= \prod_{k=0}^{K-1} \sigma_k^2$$

"eigenvalues are the diagonal entries of the diagonal matrix"

$$\underline{\underline{\Sigma}}^{-1} = \begin{bmatrix} 1/\sigma_0^2 & & 0 \\ & 1/\sigma_1^2 & \\ 0 & & \ddots \end{bmatrix}$$

$$\rightsquigarrow \mathcal{N}(\underline{x}; \underline{\mu}, \text{diag}(\underline{\sigma}^2)) = \frac{1}{\sqrt{\prod_{k=0}^{K-1}(\sigma_k^2) \cdot (2\pi)^K}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^T \begin{bmatrix} 1/\sigma_0^2 & & \\ & 1/\sigma_1^2 & \\ & & \ddots \end{bmatrix} (\underline{x}-\underline{\mu})\right)$$

in TFP: MultivariateNormalDiag

↳ but uses the scale diagonal (std per dimension instead of variance)

$$\underline{\sigma} = \sqrt{\underline{\sigma}^2} = \text{diag}(\underline{\underline{L}})$$

Spherical / Isotropical Gaussian

one variance / standard deviation for all of the dimensions

$$\underline{\underline{\Sigma}} = \sigma^2 \underline{\underline{I}} \in \mathbb{R}^{K \times K} \qquad (\underline{\underline{L}} = \sigma \underline{\underline{I}})$$

$$\det(\underline{\underline{\Sigma}}) = \prod_{k=0}^{K-1} \sigma^2 = (\sigma^2)^K = \sigma^{2K}$$

$$\underline{\underline{\Sigma}}^{-1} = (\sigma^2 \underline{\underline{I}})^{-1} = \frac{1}{\sigma^2} \underline{\underline{I}}$$

$$\mathcal{N}(\underline{x}; \underline{\mu}, \sigma^2 \underline{\underline{I}}) = \frac{1}{\sqrt{\sigma^{2K} (2\pi)^K}} \exp\left(-\frac{1}{2\sigma^2} \underbrace{(\underline{x}-\underline{\mu})^T(\underline{x}-\underline{\mu})}_{\text{just a regular scalar product}}\right)$$

TFP

- MultivariateNormalDiag with diagonal scale multiplier
- MultivariateNormalDiag with $\sigma \cdot$ tf.ones(K)
- Independent of Batched Univariate Normal