Problem 1: Simple pushforward

The Jacobian matrix $J_f(x)$ of a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ at a point x is defined as the 2×2 matrix $\left(J_f(x)\right)_{i,j} = \frac{\partial f_i}{\partial x_j}$. For the given distributions $\mathbb{P}$ and $\mathbb{Q}$, find a map f and its inverse f^{-1} such that $\mathbb{Q} = f_\# \mathbb{P}$. Additionally, compute $J_f(x)$ and $\det J_f(x)$.

1. **Shift:** $\mathbb{P}$ is $\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$, $\mathbb{Q}$ is $\mathcal{N}\left(\begin{bmatrix}1\\-2\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$;
2. **Scaling:** $\mathbb{P}$ is $\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$, $\mathbb{Q}$ is $\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}3 & 0\\0 & 0.5\end{bmatrix}\right)$;
3. **Rotation:** $\mathbb{P}$ is $\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}3 & 0\\0 & 0.5\end{bmatrix}\right)$, $\mathbb{Q}$ is $\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1.75 & 1.25\\1.25 & 1.75\end{bmatrix}\right)$;

(For this problem, you may wish to review how multivariate Gaussians transform under linear changes of variables.)

Problem 2: Infinitesimal change of variables

This problem considers a map $f : \mathbb{R}^d \to \mathbb{R}^d$ of the form $f(x) = x + u(x)\Delta$, where $u : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field and $\Delta > 0$. This map represents transporting points for a time Δ along the direction given by the vector field. Such maps are important in dynamics-based generative models, which will be discussed at the end of the course.

1. Convince yourself **informally** that for smooth enough u and small enough Δ, f is invertible.
2. Show using the variable change formula that the following identity holds for a distribution with density p and a diffeomorphism $f : \mathbb{R}^d \to \mathbb{R}^d$:
$$\log(f_\# p)(f(x)) - \log p(x) = -\log \det J_f(x).$$
3. Define $p_\Delta := f_\# p$, that is, the distribution that sampled by drawing samples x from p and transporting them for time Δ in the direction $u(x)$.
   (a) Compute $J_f(x)$ in terms of $J_u(x)$.
   (b) Show that $\log \det J_f(x) = \mathrm{Tr}\left[J_u(x)\right]\Delta + O(\Delta^2)$.
   (For this problem, you will need to have some knowledge of the matrix exponential. If this is unfamiliar to you, solve this exercise *assuming that $J_u(x)$ is a diagonal matrix*, using the Taylor expansion $\log(1+t) = t + O(t^2)$. Then read about the matrix exponential and use the fact that $\log \det A = \mathrm{Tr}\left[\log A\right]$ to solve the problem in the general case.)
   (c) Combine Part 2 of this problem with Parts 3(a,b) to show the following identity:
$$\log p_\Delta(f(x)) - \log p(x) = -\mathrm{Tr}\left[J_u(x)\right]\Delta + O(\Delta^2).$$

Observe that $\mathrm{Tr}\left[J_u(x)\right]$ is the divergence of the vector field u at x. We have thus related the change of density when transporting points along u to the divergence of u, commonly given as a physical interpretation for the divergence and an intuition for the divergence theorem in vector calculus.