Density Transformations & random number generation

5.1 Inverse CDF and Random Number Generation (4 points)

Background: If $F_X(x)$ is the cumulative distribution function (cdf) of a random variable X, then the random variable $Z = F_X(X)$ is uniformly distributed on the interval $[0, 1]$. This result provides a general receipe to generate samples $\tilde{x}$ of a random variable X with a desired probability density function (pdf) $p_X(x)$ from uniformly distributed random numbers $\tilde{z} \in [0, 1]$:

1. Compute the cdf $F_X(x)$ of the desired pdf $p_X(x)$
2. Determine the inverse transformation F^{-1}.
3. Sample uniformly distributed numbers (in $[0, 1]$), $\tilde{z}$.
4. Get the samples $\tilde{x} = F^{-1}(\tilde{z})$ from X.

The pdf of a Laplace distribution with location parameter μ (= mean), and scale parameter $b > 0$ (variance = $2b^2$) is given by

$$p_X(x) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right).$$

Task:

(a) Following the procedure above, derive a formula to generate samples of a scalar random variable with a Laplacian distribution from uniformly distributed random numbers.

(b) Implement your procedure for verification and generate 500 samples for a Laplacian random variable X with a specific mean $\mu = 1$ and scale parameter $b = 2$. Plot a density estimate (e.g. histogram, ecdf) for these samples overlayed with the pdf $p_X(x)$ from above.

5.2 Density Transformations (6 points)

Background: Let $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ be a function of $\mathbf{x} \in \Omega \subset \mathbb{R}^n$ and assume we make a change of variables to a new coordinate system by a mapping $\mathbf{u} = \mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), \ldots, u_n(\mathbf{x}))$, whose inverse mapping $\mathbf{x} = \mathbf{x}(\mathbf{u}) = (x_1(\mathbf{u}), \ldots, x_n(\mathbf{u}))$ exists and is differentiable. As we change the coordinate system, the integral over f changes according to

$$\int_\Omega f(\mathbf{x})\mathbf{dx} = \int_{u(\Omega)} f(\mathbf{x}(\mathbf{u})) \left|\det \frac{\partial \mathbf{x}(\mathbf{u})}{\partial \mathbf{u}}\right| \mathbf{du} = \int_{u(\Omega)} f(\mathbf{x}(\mathbf{u})) \frac{1}{\left|\det \frac{\partial \mathbf{u}(\mathbf{x})}{\partial \mathbf{x}}\right|}\mathbf{du},$$

where $\frac{\partial \mathbf{x}(\mathbf{u})}{\partial \mathbf{u}}$ is the Jacobi matrix, which is the matrix of the partial derivatives

$$\frac{\partial \mathbf{x}(\mathbf{u})}{\partial \mathbf{u}} = \begin{pmatrix} \frac{\partial x_1(\mathbf{u})}{\partial u_1} & \cdots & \frac{\partial x_1(\mathbf{u})}{\partial u_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n(\mathbf{u})}{\partial u_1} & \cdots & \frac{\partial x_n(\mathbf{u})}{\partial u_n} \end{pmatrix}$$

and whose determinant $\det \frac{\partial \mathbf{x}(\mathbf{u})}{\partial \mathbf{u}} = (\det \frac{\partial \mathbf{u}(\mathbf{x})}{\partial \mathbf{x}})^{-1}$ is called the *Jacobi determinant* (also *functional determinant*).

Remark: The absolute value of the Jacobi determinant at a point $\mathbf{u}_0$ corresponds to the factor by which the function $\mathbf{x}(\mathbf{u})$ expands or shrinks volumes near $\mathbf{u}_0$.

Implication: If $f(\mathbf{x})$ is the probability density function (pdf) of the n-dimensional random vector $\mathbf{X}$ then $f(\mathbf{x}(\mathbf{u})) \left|\det \frac{\partial \mathbf{x}(\mathbf{u})}{\partial \mathbf{u}}\right|$ is the pdf of the random vector $\mathbf{u}(\mathbf{X})$.

Task:

(a) Consider the density of a random variable X to be $p_X(x) = e^{-x},\ x \geq 0$. For the change of variables $u = u(x) = e^{-x}$ calculate the density $p_{u(X)}(u)$ of the random variable $u(X)$.

(b) Consider two independent and uniformly in the interval $[0, 1]$ distributed random variables $(X_1, X_2)^T =: \mathbf{X}$. The pdf is given by $p_\mathbf{X}(x_1, x_2) = 1$ in $[0, 1]^2$ and zero otherwise. Consider the variable transformation $\mathbf{u} = \mathbf{u}(\mathbf{x})$ with $u_1(\mathbf{x}) = \sqrt{-2\log x_1}\cos(2\pi x_2)$ and $u_2(\mathbf{x}) = \sqrt{-2\log x_1}\sin(2\pi x_2)$. Show that $\mathbf{u}(\mathbf{X})$ corresponds to two independent unit-variance zero-mean normally distributed random variables. *Remark:* This procedure to produce Gaussian samples from uniform random numbers is called the Box-Muller method.

(c) Outline how to extend the last result to n dimensions, i.e., how to generate samples from a multidimensional Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ just from uniformly distributed random numbers in $[0, 1]^n$. Use the following:

- Any symmetric positive semidefinite matrix (such as the covariance matrix $\boldsymbol{\Sigma}$) has a Cholesky decomposition $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$ (and that can be easily computed numerically).
- If $\mathbf{L}$ is a constant matrix and $\mathbf{X}$ a random vector then $\mathrm{Cov}(\mathbf{L}\mathbf{X}) = \mathbf{L}\,\mathrm{Cov}(\mathbf{X})\,\mathbf{L}^T$.
- The covariance matrix of independent unit-variance Gaussian variables is identity, i.e., $\mathrm{Cov}(\mathbf{X}) = \mathbf{I}$.

Total points: 10