Introduction to Deep Learning M2177.0043 March 28, 2020

Due on April 5, 2020

Seoul National University

Homework 1

INSTRUCTIONS

- The homework is due at 9:00am on April 5, 2020. Anything that is received after that time will be considered to be late and we do not receive late homeworks. We do however ignore your lowest homework grade.
- Homeworks need to be submitted electronically on ETL. Only PDF generated from LaTex is accepted.
- Make sure you prepare the answers to each question separately. This helps us dispatch the problems to different graders.
- Collaboration on solving the homework is allowed. Discussions are encouraged but you should think about the problems on your own.
- If you do collaborate with someone or use a book or website, you are expected to write up your solution independently. That is, close the book and all of your notes before starting to write up your solution.

1 Learning LaTex [0 points]

1. Insert a 2 by 2 table in LaTex and fill in (first name, last name, major, student id) in row major format.

2. Insert a jpeg or png image of your profile photo in Latex and center align the image.

2 Inverse transform sampling [30 points]

1. **Probability integral transform** A random variable X has a continuous pdf $f_X(x)$ and cdf $F_X(x)$. Prove that the random variable $Y = F_X(X)$ is uniformly distributed in $[0, 1]$.

2. **Inverse transform sampling** A random variable X has a continuous pdf $f_X(x)$ and cdf $F_X(x)$. Prove (without directly invoking the results above) that if the random variable U is uniformly distributed in $[0, 1]$, then $F_X^{-1}(U)$ has $F_X(x)$ as its CDF.

3. Use the results above to simulate the draw of 1 million samples from exponential distribution $f_X(x) = \lambda e^{-\lambda x}$ with $\lambda = 1.0$. Draw a figure overlaying following two plots: (1) analytical PDF of exponential distribution. (2) normalized histogram of your samples (use 500 bins). X-axis should be the value of X and the Y-axis should be the probability. Attach both the code in NumPy and the figure.

3 Optimal hedge ratio [10 points]

You own some shares of stock B and you just bought one share of stock A. How many shares of B do you need to sell to minimize the variance of total stocks you own? Assume you know the variance of each stocks σ_A^2, σ_B^2 and their correlation coefficient ρ.

4 Eigenvalues [10 points]

Find all eigenvalues of the $X \in \mathbb{R}^{p \times p}$ real matrix filled with some constant value c except the diagonal component. Concretely, the matrix X is structured as following:

```
X = c * np.ones((p,p)) + (-c + 1)*np.eye(p)
```

Homework 1

---

5 Multivariate normal [30 points]

1. Describe the transformation which takes n independent $N(0,1)$ standard normal samples $Z = [Z_1, \ldots, Z_n]^\intercal$, and generates correlated random variables that follow a n-dimensional multivariate normal distribution $X = [X_1, \ldots, X_n]^\intercal \sim N(\mu, \Sigma)$.

2. Simulate $10,000$ draws of $N\left(0, \Sigma = \begin{bmatrix} 1.0 & 0.9 \\ 0.9 & 1.0 \end{bmatrix}\right)$ using std samples drawn from $N(0,1)$. Contrast the results with $10,000$ draws using Numpy's existing implementation for multivariate sampling `np.random.multivariate_normal` function. Visualize three scatter plots side-by-side (in 1×3 subplot format, use `markersize=0.1` for better visualization) for 1) the std normal, 2) your transformed multivariate samples, and 3) multivariate samples obtained from the Numpy code. Does your result closely match Numpy's results? Fix the random seed with `np.random.seed(1337)` in your code to make the experiment deterministic.

6 Optimization with positivity constraints [20 points]

Consider the following minimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \;\; & f(\mathbf{x}) \\ \text{subject to} \;\; & x_i \geq 0, \; i = 1, \ldots, n, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function over $\mathbb{R}^n_+$. Regardless of the convexity of f, we saw in the lecture, that $\mathbf{x}^*$ is a local minimum solution to the constrained minimization problem iff

$$\nabla f(\mathbf{x}^*)^\intercal (\mathbf{x} - \mathbf{x}^*) \geq 0, \; \forall \, \mathbf{x} \in \mathcal{C} \tag{2}$$

Note, in the lecture, we also saw the general form of the optimality condition in Equation (2) is not practical as is since we would need to check the condition for all $\mathbf{x}$ in $\mathcal{C}$. Simplify the characterization of a local minimum solution $\mathbf{x}^*$ in Equation (2) knowing that $\mathcal{C} = \mathbb{R}^n_+$. Note, this simplification should avoid checking against all $\mathbf{x} \in \mathcal{C}$.