1. Linear Algebra and Calculus

Random Diagonal

Let $\mathbf{A} = \begin{bmatrix} a_{1,1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n,n} \end{bmatrix}$, where all entries of $\mathbf{A}$ are 0 except its diagonal entries $(a_{1,1}, \dots, a_{n,n})$.
Now suppose that each diagonal entry $a_{i,i}$ is drawn uniformly from the range $[-1, 1]$ (consider it random variable $A_{i,i}$), and assume that $n > 1$. Find:

- $p(A_{1,1} = 0)$
- $p(A_{n,n} > 0.5 \mid A_{n-1,n-1} \geq 0)$
- $p(\text{rank}(\mathbf{A}) < n)$
- $p(\mathbf{A} \succeq 0)$, i.e., the probability $\mathbf{A}$ is positive semidefinite (hint: if $n = 2$, what are the eigenvalues of $\mathbf{A}$?)

Matrix Operations

Let $\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

- Is $\mathbf{B}$ invertible? If so, find $\mathbf{B}^{-1}$.
- Is $\mathbf{B}$ diagonalizable? If so, find its diagonalization.

Derivatives of Activation Functions

The "sigmoid" and hyperbolic tangent functions are commonly used in neural networks to monotonically map a real-valued scalar into a finite range.

- Let $\sigma(x) = \frac{1}{1+e^{-x}}$. Write $\frac{d\sigma}{dx}$ in terms of $\sigma(x)$.
- Let $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Write $\frac{d}{dx}\tanh(x)$ in terms of σ and x.

2. Dynamic programming algorithm

Tile Collection

Consider an $n \times n$ grid. At position (i, j) in the grid, there is a reward $r_{i,j} > 0$. You want to travel from the top left tile (at $i = j = 1$) to the bottom right tile ($i = j = n$) while collecting rewards on the tiles you've visited. You can only move down or right one tile in each step (not diagonally).

Design a dynamic programming algorithm that maximizes (and outputs) the sum of your reward when you reach the bottom right tile. A 3×3 example is shown below; the optimal path is colored in red. Give the time and space complexity of your algorithm.

1	2	8
6	5	5
3	4	1