2023-August-Mathematics of Network Algorithms

Problem Set

(Last Updated: August 27, 2023)

Linear Algebra

- Define the following terms (and every term used to define them) with an illustrative example:
  (a) Linear dependence and span of vectors (b) Norm of a vector
  (c) Eigenvalue, eigenvector and eigendecomposition
- Prove that ℓ_1, ℓ_2, and ℓ_∞ norm satisfies the properties mentioned while answering 1 (b).
- Write a 3×3 matrix $\mathbf{A}$ that is *not* identity, nor symmetric nor orthogonal. Also, write $\mathbf{A} \times \mathbf{A}$ and its transpose, inverse, determinant, eigenvalues, and eigenvectors.
- Write a (non-trivial) system of linear equations with at least 4 variables and 5 constraints both in equation form and matrix form.
- Consider a function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ defined as $f((x_1, x_2)) = (-2x_2, -3x_1 + x_2)$. Describe geometrical interpretation of the above function in terms translation matrix, mention how a random point is shifted, and special points that may be only stretched.
- Consider a $n \times n$ matrix A. Prove that $||A||_F^2 = \mathrm{Tr}(A \cdot A^T)$.
- Prove that any symmetric matrix has real eigen values and the corresponding eigen vectors are orthogonal to each other.

Probability and Stastictics

- Describe the following terms with an illustrative examples:
  (a) Probability Space (b) Random variables
  (c) Frequentist probability and Bayesian probability
  (d) Mean, Variance, Covariance, & Correlation
- We flip a fair coin ten times. Find the probability of the following events: (i) Nr of heads and talks are equal. (ii) Nr of heads is more than nr of tails. (iii) The i^{th} flip and $(11-i)^{\text{th}}$ flip are same for every $i \in [5]$.
- We roll two fair dice. What is the probability space? What is the expectation of random variable representing the sum of two dice?
- Define the following distributions:
  (a) Bernoulli Distribution (b) Gaussian Distribution
  (c) Laplace Distributions (d) Multinoulli Distribution
  (e) Uniform Distribution
- Select your favourite distribution and derive expressions for its (i) expectation, (ii) variance, and (iii) standard deviation.

Numerical Optimization

- Consider the univariate function $f(x) = x^3 + 6x^2 - 3x - 5$. Find its stationary points and indicate whether they are maximum, minimum, or saddle points.
- Describe *overflow*, *underlow*, and *poor conditioning* with examples.
- Define *gradient* and *directional derivative*.
- Computer $\partial(f)/\partial\mathbf{x}$ when (i) $f = \sin(x_1)\cos(x_2)$, (ii) $f = 4x_1^2x_3 + 4x_1x^2x_3 + 5x_3^4$, and (iii) $f = x_1x_2x_4 + 2x_3^2x_4 + \sin(x_1x_2x_3)$.
- Prove the following identities:
  - $\partial(\mathbf{x}^\top\mathbf{x})/\partial\mathbf{x} = 2\mathbf{x}^\top$, $\partial(\mathbf{x}^\top\mathbf{a})/\partial\mathbf{x} = \mathbf{a}^\top$ and $\partial(\mathbf{a}^\top\mathbf{x})/\partial\mathbf{x} = \mathbf{a}^\top$
  - $\partial(\mathbf{a}^\top\mathbf{B}\mathbf{x})/\partial\mathbf{x} = \mathbf{B}^\top\mathbf{a}$, and $\partial(\mathbf{x}^\top\mathbf{B}\mathbf{x})/\partial\mathbf{x} = \mathbf{x}^\mathsf{T}(\mathbf{B} + \mathbf{B}^\top)$
  - For symmetric matrix W, $\partial((\mathbf{x} - \mathbf{A}\mathbf{s})^\top\mathbf{W}(\mathbf{x} - \mathbf{A}\mathbf{s}))/\partial\mathbf{s} = -2(\mathbf{x} - \mathbf{A}\mathbf{s})^\top\mathbf{W}\mathbf{A}$.
- Prove that a function $f(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{R}$ decreases fastest in the direction opposite to its gradient.
- Consider the optimization problem $\min\{\frac{1}{2}\mathbf{w}^\top\mathbf{w}\}$ over all $\mathbf{w} \in \mathbb{R}^n$ subjected to $\mathbf{w}^\top\mathbf{w} \geq 1$. Convert it into an unconstrained optimization problem by introducing Lagrange multiplier λ.

Machine Learning

- Describe the following terms with an illustrative examples:
  (a) Artificial Intelligence
  (b) Machine Learning
  (c) Deep Learning
  (d) Perceptron
  (e) Neural Network
  (f) Activation function
  (g) Loss Function
  (h) Optimisers
  (i) Parameters and Hyperparameters
- Write steps in Principal Component Analysis to reduce 2-dimension data to 1-dimension data.
- Define *learning* in the context of Machine Learning.
- Write short description on five types of *tasks* (in the context of Machine Learning).
- What is *supervised learning* and *unsupervised learning*?
- Consider a learner regression problem where the objective is determine the value of $\mathbf{w} \in \mathbb{R}^n$ such that $\mathbf{w}^\top\mathbf{x}$ is as close to y as possible for vector $\mathbf{x}_i \in \mathbb{R}^n$ and scalar y_i for all $i \in [m]$. Derive an analytical expression to compute w if the difference between actual and computed values is determined using mean squared error.