CM146, Winter 2019
Problem Set 0: Math prerequisites
Due January 13, 2020

Jacob Kaufman

01/13/2019

1 Multivariate Calculus

Consider $y = z\sin(x)e^{-x}$. What is the partial derivative of y with respect to x?

Solution:

$$y = z\sin(x)e^{x}$$
$$\frac{\partial y}{\partial x} = z \cdot \frac{\partial}{\partial x}(\sin(x)e^{x})$$
$$\frac{\partial y}{\partial x} = z \cdot (\sin(x) \cdot \frac{\partial}{\partial x}e^{-x} + \frac{\partial}{\partial x}\sin(x) \cdot e^{-x})$$
$$\frac{\partial y}{\partial x} = z(\sin(x)(-e^{-x}) + \cos(x)e^{-x})$$
$$\frac{\partial y}{\partial x} = ze^{-x}(\cos(x) - \sin(x))$$

2 Linear Algebra

Consider the matrix $\boldsymbol{X}$ and the vectors $\boldsymbol{y}$ and $\boldsymbol{z}$ below:

$$\boldsymbol{X} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \boldsymbol{z} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

(a) What is the inner product $\boldsymbol{y}^T\boldsymbol{z}$?
Solution:

$$\boldsymbol{y}^T\boldsymbol{z} = \begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 + 6 = 8$$

(b) What is the product $\boldsymbol{X}\boldsymbol{y}$?
Solution:

$$\boldsymbol{X}\boldsymbol{y} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 + 12 \\ 1 + 6 \end{pmatrix} = \begin{pmatrix} 14 \\ 7 \end{pmatrix}$$

(c) Is $\boldsymbol{X}$ invertible? If so, give the inverse; if not, explain why not.
Solution:

No, $\boldsymbol{X}$ is not invertible because its columns are linearly dependent.

(d) What is the rank of $\boldsymbol{X}$**?**
Solution:

$$\textbf{rank}(\boldsymbol{X}) = \textbf{rank} \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} = \textbf{rank} \begin{pmatrix} 2 & 4 \\ 0 & 0 \end{pmatrix} = \textbf{rank} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = 1$$

3 Probability and Statistics

(a) What is the sample mean for this data?
Solution:
$$\mu = \frac{1+1+1}{5} = \frac{3}{5}$$

(b) What is the unbiased sample variance?
Solution:
$$\sigma^2 = \frac{(1-\frac{3}{5})^2 + (1-\frac{3}{5})^2 + (1-\frac{3}{5})^2 + (0-\frac{3}{5})^2 + (0-\frac{3}{5})^2}{5} = 0.24$$

(c) What is the probability of observing this data assuming that a coin with an equal probability of heads and tails was used?
Solution: $P = (0.5)^5 = 0.03125$

(d) What is the value that maximizes the probability of the sample S?
Solution:
$$\begin{aligned}
P &= p \cdot p \cdot (1-p) \cdot p \cdot (1-p) && (1)\\
&= p^3(1-p)^2 && (2)\\
&= p^3(1-2p+p)^2 && (3)\\
&= p^3 - 2p^4 + p^5 && (4)\\
\frac{dP}{dp} &= 5p^4 - 8p^3 + 3p^2 = 0 && (5)\\
5p^2 - 8p + 3 &= 0 && (6)\\
(5p-3)(p-1) &= 0 && (7)\\
p &= \frac{3}{5} = P(X_i = 1) && (8)
\end{aligned}$$

(e) Given the following joint distribution between X and Y , what is $P(X = T \mid Y = b)$?
Solution: $P(X = T \mid Y = b) = \frac{0.1}{0.1+0.5} = 0.4$

4 Probability Axioms

(a) **Solution:** False

(b) **Solution:** False except when $P(A \cap B) = 0$

(c) **Solution:** False

(d) **Solution:** False

(e) **Solution:** False

5 Discrete and Continuous Distributions

(a) **Solution:** (v)

(b) **Solution:** (iv)

(c) **Solution:** (ii)

(d) **Solution:** (i)

(e) **Solution:** (iii)

6 Mean and Variance

(a) What is the mean and variance of a $Bernoulli(p)$ random variable?
Solution:
$\mathbf{X} \sim Bernoulli(p)$
$\mathrm{E}(\mathbf{X}) = p$
$\mathrm{Var}(\mathbf{X}) = p(1-p)$

(b) If the variance of a zero-mean random variable X is σ^2, what is the variance of $2X$? What about the variance of $X+3$?
Solution:

$$Var(aX) = a^2 Var(X) \implies Var(2X) = 4Var(X) = 4\sigma^2$$

$$Var(X+b) = Var(X) \implies Var(X+3) = Var(X) = \sigma^2$$

7 Algorithms

(a) **Big-O Notation**

i. $f(n) = ln(n),\ g(n) = lg(n)$
Solution:

$$ln(n) = \frac{lg(n)}{lg(e)} \tag{9}$$

$$k = lg(e) \tag{10}$$

$$ln(n) = \frac{1}{k} lg(n) \tag{11}$$

Thus both are correct. $f(n) = O(g(n))$ and $g(n) = O(f(n))$

ii. $f(n) = 3^n, g(n) = n^{10}$
Solution:

$$\lim_{x\to\infty} \frac{n^{10}}{3^n} = 0 \implies g(n) = O(f(n))$$

iii. $f(n) = 3^n, g(n) = 2^n$
Solution:

$$\lim_{x\to\infty} \frac{2^n}{3^n} = 0 \implies g(n) = O(f(n))$$

(b) Divide and Conquer
Solution: Start by assigning a left pointer to the first index in the array and a right pointer to the last index in the array. Recursively, consider the midpoint between the two pointers. If the midpoint is the index of the left pointer, return that index. Else, if the midpoint has value zero, move the left pointer to the midpoint. Else move the right pointer to the midpoint. This is correct because once we confirm the value of a midpoint, we can move a pointer with the same value to the midpoint without losing information, as we know the elements we stop considering will be all 1s or all 0s. This runs in O(log(n)) time because the length of the list we consider is halved by each iteration.

8 Probability and Random Variables

(a) If X and Y are independent random variables, show that $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.
Solution:

$$\mathbb{E}[XY] = \sum_i \sum_j f_{xy}(x_i y_j) \tag{12}$$

$$= \sum_i \sum_j x_i y_j f_x(x_i) f_y(y_j) \tag{13}$$

$$= (\sum_i x_i f_x(x_i))(\sum_j y_j f_y(y_j)) \tag{14}$$

$$= \mathbb{E}[X]\mathbb{E}[Y] \tag{15}$$

(b) Provide one line justifications

i. **Solution:** The law of large numbers states that the increasing number of samples will approach the population distribution, which is expected to be $\frac{1}{6}$ the sample size in the case of a fair die.

ii. **Solution:** The Central Limit Theorem states that

$$\sqrt{n}(X - \mu) \xrightarrow{n \to \infty} \mathcal{N}(0, \sigma^2)$$

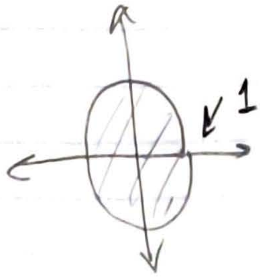
, so with a Bernoulli distribution of $\mu = \frac{1}{2}, Var(X) = \sigma^2 = \frac{1}{4}$, $\sqrt{n}(X - \frac{1}{2}) \xrightarrow{n \to \infty} \mathcal{N}(0, \frac{1}{4})$.

9 Linear Algebra

(a) **Vector Norms**

Solution: The hand-drawn solutions appear on the next page.

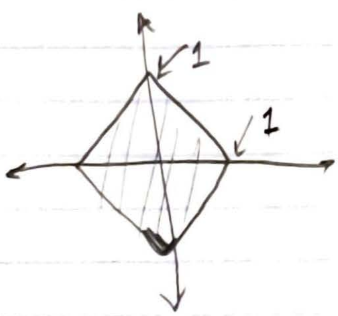
i.)



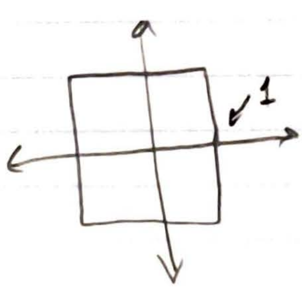
ii.)

Any points on the axes

iii.)



iv)



(b) **Matrix Decompositions**

i. Give the definitions of the eigenvalues and eigenvectors of a square matrix.

Solution: For a square matrix $\mathbf{A}$, the eigenvalue λ, and corresponding eigenvector $\mathbf{v}$ satisfy

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

ii. Find the eigenvalues and eigenvectors of

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Solution:

Finding the eigenvalues:

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$$

$$\det\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = 0$$

$$(2-\lambda)^2 - 1 = 0$$

$$\lambda^2 - 4\lambda + 4 - 1 = 0$$

$$\lambda^2 - 4\lambda + 3 = 0$$

$$(\lambda - 3)(\lambda - 1) = 0$$

$$\lambda = 3, \lambda = 1$$

Finding the corresponding eigenvectors:

$$(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$$

For $\lambda = 1$,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0} \implies \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} t \\ -t \end{pmatrix}$$

For $\lambda = 3$,

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\mathbf{v} = \mathbf{0} \implies \mathbf{v} = \begin{pmatrix} t \\ t \end{pmatrix}$$

Eigenvalue, eigenvector pairs:

$$\lambda = 1, \mathbf{v} = \begin{pmatrix} t \\ -t \end{pmatrix}$$

$$\lambda = 3, \mathbf{v} = \begin{pmatrix} t \\ t \end{pmatrix}$$

iii. For any positive integer k, show that the eigenvalues of $\boldsymbol{A}^k$ are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$, the kth powers of the eigenvalues of matrix $\boldsymbol{A}$, and that each eigenvector of $\boldsymbol{A}$ is still an eigenvector of $\boldsymbol{A}^k$.

Solution:

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}$$

$$\boldsymbol{A}^2\mathbf{v} = \lambda(\lambda\mathbf{v})$$

$$\vdots$$

$$\boldsymbol{A}^k\mathbf{v} = \lambda(\boldsymbol{A}^{k-1})\mathbf{v} = \lambda \cdot \lambda^{k-1}\mathbf{v}$$

$$\implies \boldsymbol{A}^k\mathbf{v} = \lambda^k\mathbf{v}$$

(c) Vector and Matrix Calculus

i. What is the first derivative of $\boldsymbol{a}^T\boldsymbol{x}$ with respect to $\boldsymbol{x}$?

Solution:

$$\frac{d}{d\boldsymbol{x}}(\boldsymbol{a}^T\boldsymbol{x}) = \frac{d}{d\boldsymbol{x}} a_1x_1 + \cdots + a_nx_n = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{a}$$

ii. What is the first derivative of $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ with respect to $\boldsymbol{x}$? What is the second derivative?

Solution:

$$\frac{d}{d\boldsymbol{x}}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x} \xrightarrow{symmetric} 2\boldsymbol{A}\boldsymbol{x}$$

$$\frac{d^2}{d\boldsymbol{x}^2}(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}) = \frac{d}{d\boldsymbol{x}}(2\boldsymbol{A}\boldsymbol{x}) = 2\boldsymbol{A}$$

(d) Geometry

i. Show that the vector $\boldsymbol{w}$ is orthogonal to the line $\boldsymbol{w}^T\boldsymbol{x} + b = 0$.
Solution:
Consider $\mathbf{x}_1, \mathbf{x}_2$ on the line $\boldsymbol{w}^T\boldsymbol{x}+b = 0$, where $\boldsymbol{x}_1 - \boldsymbol{x}_2$ represents an arbitrary vector in the span of the line.

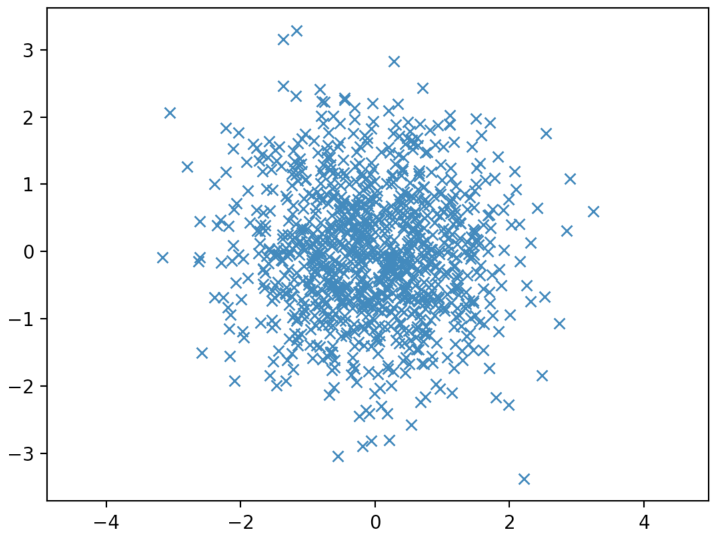
$$\boldsymbol{w}_T(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \boldsymbol{w}^T\boldsymbol{x}_1 - \boldsymbol{w}^T\boldsymbol{x}_2 = b - b = 0$$

Because the inner product between $\boldsymbol{w}$ and $\boldsymbol{x}_1 - \boldsymbol{x}_2$ is 0.

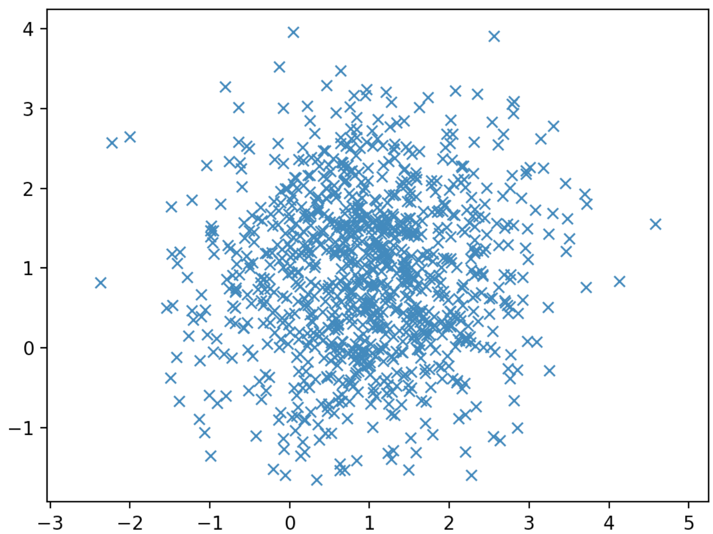
ii. Argue that the distance from the origin to the line $\boldsymbol{w}^T\boldsymbol{x} + b = 0$ is $\frac{b}{||x||_2}$.
Solution: The length of any vector from the origin to a point on this line is equal to $\frac{b}{length}$.

10 Sampling from a Distribution

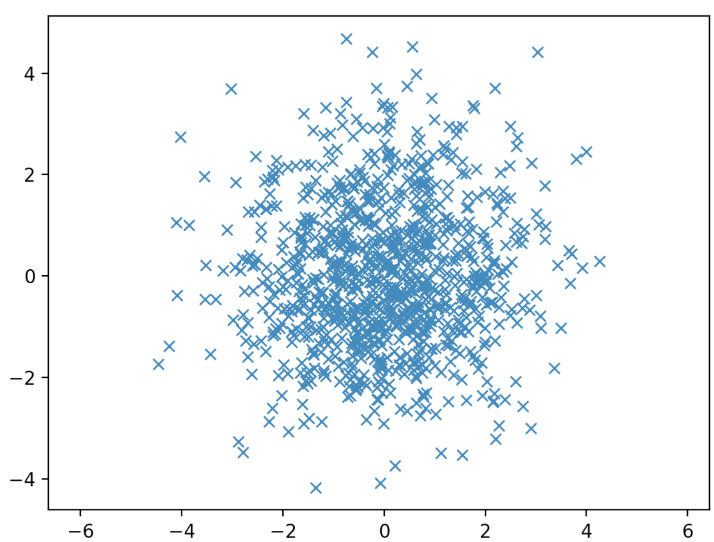
(a) Plot 1



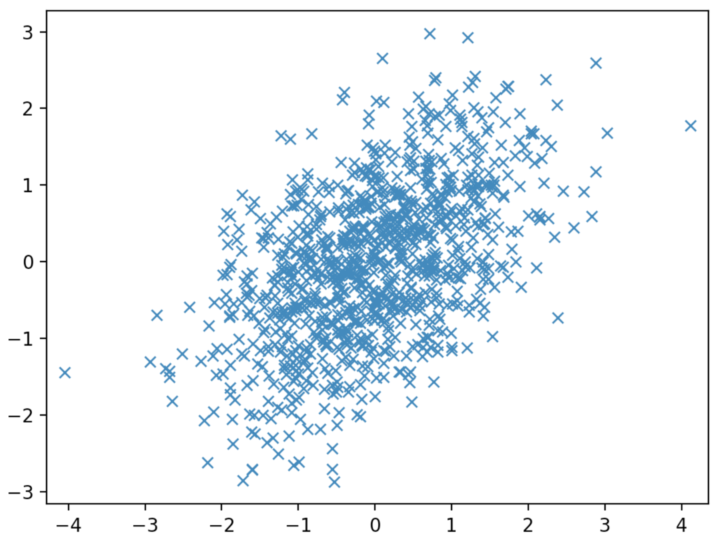
(b) Plot 2



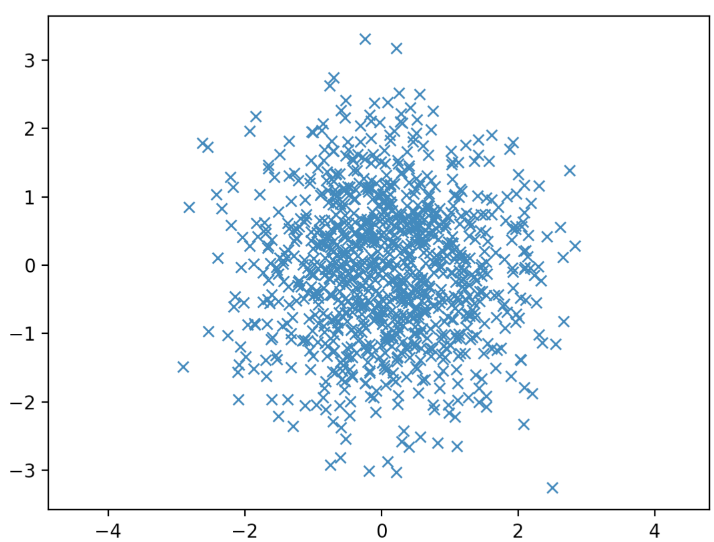
(c) Plot 3



(d) Plot 4



(e) Plot 5



11 Eigendecomposition

Write a python program to compute the eigenvector corresponding to the largest eigenvalue of the following matrix and submit the computed eigenvector.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix}$$

Solution:

$$\mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \lambda = 3$$

.

12 Data

(a) ”11k Hands”

(b) sites.google.com/view/11khands

(c) This data set contains 11,076 images of hands. It uses data such as whether the image is of the left or right hand, as well as indicators of whether the hand has irregularities, accessories, or nail polish to predict gender and biometrically identify the subject.

(d) 11,076

(e) 4