COT 5615 Math for Intelligent Systems Fall 2021 Homework #2

UFID: 96703101
Name: *Vyom Pathak*
Instructor: Professor Kejun Huang
Due Date: September 13, 2021

Problem 3.2

RMS value and average of block vectors

Solution

Vector x can be shown as follows: $x = (a_1, \ldots, a_n, b_1, \ldots, b_m)$

(a) RMS of vector x is:

$$\begin{aligned} rms(x) &= \sqrt{\frac{a_1^2 + \ldots + a_n^2 + b_1^2 + \ldots + b_m^2}{n+m}} \\ &= \sqrt{\frac{\frac{n(a_1^2+\ldots+a_n^2)}{n} + \frac{m(b_1^2+\ldots+b_m^2)}{m}}{n+m}} \\ &= \sqrt{\frac{n\ rms(a)^2 + m\ rms(b)^2}{n+m}} \end{aligned}$$

(b) The average of vector x is:

$$\begin{aligned} avg(x) &= \frac{a_1 + \ldots + a_n + b_1 + \ldots + b_m}{n+m} \\ &= \frac{\frac{n(a_1+\ldots+a_n)}{n} + \frac{m(b_1+\ldots+b_m)}{m}}{n+m} \\ &= \frac{n\ avg(a) + m\ avg(b)}{n+m} \end{aligned}$$

Problem 3.5

General norms

Solution

For 1-norm

(a) Homogeneity:

$$\begin{aligned} \|\beta x\|_1 &= |\beta x_1| + |\beta x_1| + \ldots + |\beta x_n| \\ &= |\beta|\ |x_1| + |\beta|\ |x_1| + \ldots + |\beta|\ |x_n| \\ &= |\beta|(|x_1| + |x_1| + \ldots + |x_n|) \\ &= |\beta|\ \|x\|_1 \end{aligned}$$

(b) Triangle inequality:

$$\begin{aligned} \|x+y\|_1 &= |x_1 + y_1| + |x_2 + y_2| + \ldots + |x_n + y_n| \\ &\leq |x_1| + |x_2| + \ldots + |x_n| + |y_1| + |y_2| + \ldots + |y_n| \\ &= \|x\|_1 + \|y\|_1 \end{aligned}$$

(c) Nonnegativity: Each term in $\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n|$ is nonnegative.

(d) Definiteness: $\|x\|_1 = 0$ only if $|x_1| = \ldots = |x_n| = 0$

For ∞-norm

(a) Homogeneity:

$$\begin{aligned}\|\beta x\|_\infty &= max\{|\beta x_1|, |\beta x_2|, \ldots, |\beta x_n|\} \\ &= max\{|\beta|\ |x_1|, |\beta|\ |x_2|, \ldots, |\beta|\ |x_n|\} \\ &= |\beta| max\{|x_1|, |x_2|, \ldots, |x_n|\} \\ &= |\beta|\ \|x\|_\infty\end{aligned}$$

(b) Triangle inequality:

$$\begin{aligned}\|x + y\|_\infty &= max\{|x_1 + y_1|, |x_2 + y_2|, \ldots, |x_n + y_n|\} \\ &\leq max\{|x_1| + |y_1|, |x_2| + |y_2|, \ldots, |x_n| + |y_n|\} \\ &\leq max\{|x_1|, |x_2|, \ldots, |x_n|\} + max\{|y_1|, |y_2|, \ldots, |y_n|\} \\ &= \|x\|_\infty + \|y\|_\infty\end{aligned}$$

(c) Nonnegativity: According to it's definition, $\|x\|_\infty$ is the largest of n nonnegative numbers $|x_j|$.

(d) Definiteness: $\|x\|_\infty = 0$ only if $|x_j| = 0$ for $j = 1, \ldots, n$

Problem 3.16

Effect of scaling and offset on average and standard deviation

Solution

(a) Finding $avg(\alpha x + \beta)$ as follows:

$$\begin{aligned}avg(\alpha x + \beta) &= \frac{1^T(\alpha x + \beta 1)}{n} \\ &= \frac{\alpha 1^T x + \beta 1^T 1}{n} \\ &= \alpha avg(x) + \beta \quad (\because 1^T 1 = n)\end{aligned}$$

(b) Finding $std(\alpha x + \beta)$ using (a) as follows:

$$\begin{aligned}std(\alpha x + \beta) &= rms(\alpha x + \beta - (\alpha avg(x) + \beta)1) \\ &= rms(\alpha x - \alpha avg(x)1) \\ &= |\alpha| rms(x - avg(x)1) \\ &= |\alpha| std(x)\end{aligned}$$

Problem 3.26

Time series auto-correlation

Solution

(a) R(0) is the correlation coefficient between x and x with no lag and thus, it is always 1. For R(τ), with de-mean values as follows $(0_\tau, x - \mu 1)$ and $(x - \mu 1, 0_\tau)$ and having the same norm $\|x - \mu 1\|$; their

correlation coefficient is calculated as follows:

$$R(\tau) = \frac{(0_\tau, x - \mu 1^T)(x - \mu 1, 0_\tau)}{\|x - \mu 1\|^2}$$

For $\tau \geq T$, the inner product in the numerator is 0 as for each jth vector-pair, one of the two vector has a 0 entry.

(b) We can express $R(\tau)$ as follows:

$$\begin{aligned} R(\tau) &= \left(\frac{(0_\tau, x - \mu 1)}{\|x - \mu 1\|}\right)^T \left(\frac{(x - \mu 1, 0_\tau)}{\|x - \mu 1\|}\right) \\ &= \left(\frac{(0_\tau, x - \mu 1)}{\sqrt{T} std(x)}\right)^T \left(\frac{(x - \mu 1, 0_\tau)}{\sqrt{T} std(x)}\right) \\ &= \frac{1}{T}(0_\tau, z)^T (z, 0_\tau) \end{aligned}$$

Here, according to (a) $R(\tau) = 0$ for $\tau \geq T$ and thus the above equation for the remaining indices is given as follows:

$$R(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} z_t z_{t+\tau}$$

(c) The series $x = (+1, -1, +1, -1, \ldots, +1, -1)$ has mean zero and norm $\sqrt{T}$. The auto-correlation can be calculated as follows:

$$\begin{aligned} R(\tau) &= \frac{1}{T} \sum_{t=1}^{T-\tau} x_t x_{t+\tau} \\ &= \frac{1}{T} \sum_{t=1}^{T-\tau} (-1)^{t+1} (-1)^{t+\tau+1} \\ &= \frac{T - \tau}{T} (-1)^\tau \end{aligned}$$

(d) R(7) being high means that z_t and z_{t+7} are either higher than or lower than the mean value together. For examples, as x_{t+7} is the number of meals served 7 days after x_t; number meals on Mondays are higher than the mean value and the number of meals are lower than the mean value on Saturdays. (Considering week starting from Saturday).

Problem A3.2

Nearest neighbor and smallest angle

Solution

```
using LinearAlgebra
a=[1,3,4]
xn = [[4,3,5],[0.4,10,50],[1,4,10],[30,40,50]]
dist = []
angle = []
for i in 1:4
    ntp = norm(a-xn[i])
```

```
    # println("Distance of a from x£i £ntp")
    push!(dist, ntp)
end
println("Minimum distance from a to x1,...,x4 is ",string(minimum(dist)))
for i in 1:4
    atp = acos((a'xn[i])/(norm(a)*norm(xn[i])))
    # println("Angle between a and x£i £atp")
    push!(angle, atp)
end
println("Minimum angle between a and x1,...,x4 is ",string(minimum(angle)))
```

Problem A3.5

Solution

The $\|\alpha a + \beta b + \gamma c\|$ can be solved as follows:

$$\begin{aligned}\|\alpha a + \beta b + \gamma c\|^2 &= (\alpha a + \beta b + \gamma c)^T(\alpha a + \beta b + \gamma c)\\ &= \alpha^2 a^T a + \alpha\beta a^T b + \alpha\gamma a^T c + \beta\alpha b^T a + \beta^2 b^T b + \beta\gamma b^T c + \gamma\alpha c^T a + \gamma\beta c^T b + \gamma^2 c^T c\\ &= \alpha^2\|a\|^2 + \beta^2\|b\|^2 + \gamma^2\|c\|^2 \ (\because \textit{ inner product of orthogonal vectors is } 0.)\end{aligned}$$

$$\therefore \|\alpha a + \beta b + \gamma c\| = \sqrt{\alpha^2\|a\|^2 + \beta^2\|b\|^2 + \gamma^2\|c\|^2}$$

Problem A4.2

Topic discovery via k-means

Solution

(a) Following is the code:

```
include("wikipedia_corpus.jl")
include("kmeans.jl")
using Statistics
using LinearAlgebra
using SparseArrays
using Plots
pyplot()
#k=2
centroids, labels, j_hist = Kmeans.kmeans(article_histograms, 2);
centroids_2, labels_2, j_hist_2 = Kmeans.kmeans(article_histograms, 2);
itr1 = []
itr2 = []
for i in 1:length(j_hist)
    push!(itr1,i)
end
for i in 1:length(j_hist_2)
    push!(itr2,i)
end
plot(j_hist,itr1,label="Iteration 1",title="2 Centroids")
plot!(j_hist_2,itr2,label="Iteration 2")
#k=5
```

```
centroids, labels, j_hist = Kmeans.kmeans(article_histograms, 5);
centroids_2, labels_2, j_hist_2 = Kmeans.kmeans(article_histograms, 5);
itr1 = []
itr2 = []
for i in 1:length(j_hist)
    push!(itr1,i)
end
for i in 1:length(j_hist_2)
    push!(itr2,i)
end
plot(j_hist,itr1,label="Iteration 1",title="5 Centroids")
plot!(j_hist_2,itr2,label="Iteration 2")
#k=10
centroids, labels, j_hist = Kmeans.kmeans(article_histograms, 10);
centroids_2, labels_2, j_hist_2 = Kmeans.kmeans(article_histograms, 10);
itr1 = []
itr2 = []
for i in 1:length(j_hist)
    push!(itr1,i)
end
for i in 1:length(j_hist_2)
    push!(itr2,i)
end
plot(j_hist,itr1,label="Iteration 1",title="10 Centroids")
plot!(j_hist_2,itr2,label="Iteration 2")
```

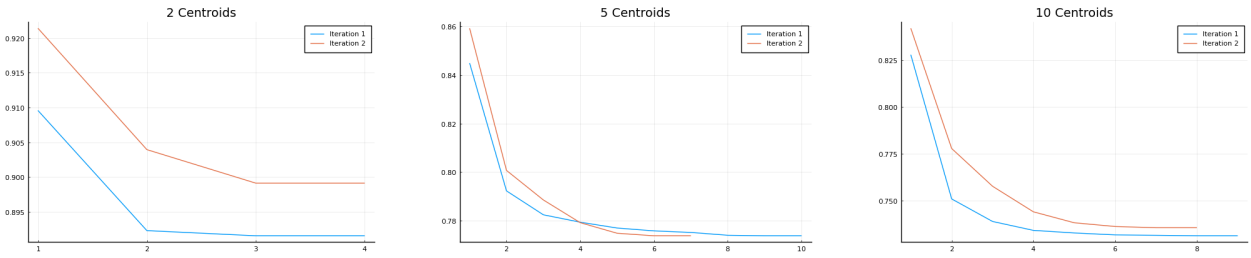

Figure 1: Loss v/s Iterations Graphs for K = 2,5,10

As I increased the number of centroids, the overall loss kept decreasing which indicates that there are at least 10 classes for the Wikipedia Article Corpus. Also, the loss is decreasing in similar manner even when the experiment is done twice for a single K-value.

(b) Table 1 describes the details learned by each topic and the three most common words for each topic as well. I used k=10 as the number of clusters. Code used to generate the data and analyze it:

```
open("Cluster_Data.txt", "w") do file
    for i in 1:10
        topic = "Topic Number: "*string(i)
        titles = article_titles[labels_2 .== i]
        words = dictionary[sortperm(centroids_2[i],rev=true)]
        ans = topic*"\nTitles: "*join(titles,",")*" \nCommon Words: "*join(words,",")*"\n\n"
        write(file, ans)
    end
end
```

Topic	3 Most Common Words	Description
1	light,solar,sun	This topic is about radiation, Sun light, and mostly related to effect of sun light and other articles related to sun. It also contains articles on different effects of sun.
2	painting,art,paintings	This topic is about different painters and different types of art and paintings. It also shows different types of painting styles in french.
3	international,convention,member	This topic is about organizations, politics and different governing bodies. It encompasses different international organizations related to finances and other socioeconomic matters.
4	signal,digital,telephone	This topic is about waves, signals and more specifically telephonic waves and its applications. It also explains wide variety of protocols related to telephonic and communication equipment.
5	nations,council,general	This topic attributes about the national and international people, food-fodder and refugee committees. . It contains articles on world food orders, world peace meets, and other organizations which maintain balance between mankind throughout the world.
6	weather,pressure,air	This topic is about effects and phenomenon related to weather, and other instruments used to measure them. It also contains articles on weather maps,and weather modifications.
7	radio,signal,frequency	This topic is about signals and phenomenons related to signals and radio waves. It also contains articles on different effects and types of waves.
8	pokemon,game,games	This topic is about Pokemon's; their types and different names and their details.
9	ice,rain,freezing	This topic is about different types of ice and ice related weather events and other ice phenomenons.
10	wind,temperature,humidity	This topic is about wind and different equipment's as well as terms related to the study of wind and temperature.

Table 1: Topic details for 10-Clusters on Wikipedia Article Corpus