Homework 3

1. Linear combinations of cash flows. We consider cash flow vectors over $T$ time periods, with a positive entry meaning a payment received, and negative meaning a payment made. A (unit) single period loan, at time period $t$, is the $T$-vector $l_t$ that corresponds to a payment received of $1$ in period $t$ and a payment made of $(1 + r)$ in period $t + 1$, with all other payments zero. Here $r > 0$ is the interest rate (over one period).

(a) Show that $l_1, \ldots, l_{T-1}$ are linearly independent.

(b) Let $c$ be a $1\ T-1$ period loan, starting at period 1. This means that $1$ is received in period 1, $(1 + r)^{T-1}$ is paid in period $T$, and all other payments (i.e., $c_2, \ldots, c_{T-1}$) are zero. Express $c$ as a linear combination of single period loans.

2. Orthogonalizing vectors. Suppose that $a$ and $b$ are any $n$-vectors. Show that we can always find a scalar $\gamma$ so that $(a - \gamma b) \perp b$. (Give a formula for the scalar $\gamma$.) Roughly speaking, we can always subtract a multiple of a vector from another one, so that the result is orthogonal to the original vector. This is called orthogonalization, and is a basic idea used in the Gram-Schmidt algorithm.

3. Vandermonde matrices. A Vandermonde matrix is an $m \times n$ matrix of the form

$$V = \begin{bmatrix}
1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\
1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & t_m & t_m^2 & \cdots & t_m^{n-1}
\end{bmatrix}$$

where $t_1, \ldots, t_m$ are numbers. We will assume that these numbers are distinct, i.e., different from each other. Multiplying the Vandermonde matrix $V$ by an $n$-vector $c$ is the same as evaluating the polynomial of degree less than $n$, with coefficients $c_1, \ldots, c_n$, at the points $t_1, \ldots, t_m$; see page 86 in the book.

Show that the columns of a Vandermonde matrix are linearly independent provided $m \geq n$.

**Hint.** You can use the following fact from algebra: If a polynomial $p$ with degree less than $n$ has $n$ or more roots (points $t$ for which $p(t) = 0$) then all its coefficients are zero.

4. Vandermonde matrices in Julia. Write a function that takes a positive integer $n$ and an $m$-vector $t$ as inputs and generates the corresponding $m \times n$ Vandermonde matrix, as described in the previous problem.
5. **Student group membership.** Let \( G \in \mathbb{R}^{m \times n} \) represent a contingency matrix of \( m \) students who are members of \( n \) groups:

\[
G_{ij} = \begin{cases} 
1 & \text{student } i \text{ is in group } j \\
0 & \text{student } i \text{ is not in group } j.
\end{cases}
\]

(A student can be in any number of the groups.)

(a) What is the meaning of the 3rd column of \( G \)?

(b) What is the meaning of the 15th row of \( G \)?

(c) Give a simple formula (using matrices, vectors, etc.) for the \( n \)-vector \( M \), where \( M_i \) is the total membership (i.e., number of students) in group \( i \).

(d) Interpret \((GG^T)_{ij}\) in simple English.

(e) Interpret \((G^T G)_{ij}\) in simple English.

6. **Linear functions.** For each description of \( y \) below, express it as \( y = Ax \) for some \( A \). (You should specify \( A \).)

(a) \( y_i \) is the difference between \( x_i \) and the average value of \( x_1, \ldots, x_{i-1} \). (We take \( y_1 = x_1 \).)

(b) \( y_i \) is the difference between \( x_i \) and the average value of all other \( x_j \)'s, i.e., the average value of \( x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \).

7. **Matrix power identity.** A student says that for any square matrix \( A \),

\[
(A + I)^3 = A^3 + 3A^2 + 3A + I.
\]

Is she right? If she is, explain why; if she is wrong, give a specific counterexample, i.e., a square matrix \( A \) for which it does not hold.

8. **Topic discovery via k-means.** In this problem you will use k-means to cluster 300 Wikipedia articles selected from 5 broad groups of topics. The Julia file `wikipedia_corpus.jl` contains the histograms as a list of 300 1000-vectors in the variable `article_histograms`. It also provides the list of article titles in `article_titles` and a list of the 1000 words used to create the histograms in `dictionary`.

The file `kmeans.jl` provides a Julia implementation of the k-means algorithm in the function `kmeans`. The `kmeans` function accepts a list of vectors to cluster along with the number of clusters, \( k \), and returns three things: the centroids as a list of vectors, a list containing the index of each vector’s closest centroid, and a list of the value of \( J \) after each iteration of k-means. Each time the function `kmeans` is invoked it initializes the centroids by randomly assigning the data points to \( k \) groups and taking the \( k \) representatives as the means of the groups. (This means that if you run `kmeans` twice, with the same data, you might get different results.)

For example, here is an example of running k-means with \( k = 8 \) and finding the 30th article’s centroid.
include("wikipedia_corpus.jl")
include("kmeans.jl")
using Kmeans

centroids, labels, j_hist = kmeans(article_histograms, 8)
centroids[labels[30]]

The list \texttt{labels} contains the index of each vector’s closest centroid, so if the 30th entry in \texttt{labels} is 7, then the 30th vector’s closest centroid is the 7th entry in \texttt{centroids}.

There are many ways to explore your results. For example, you could print the titles of all articles in a cluster.

\texttt{julia> article\_titles[labels .== 7]}
16-element Array{UTF8String,1}:
  "Anemometer"
  "Black ice"
  "Freezing rain"
  ...

Alternatively, you could find a topic’s most common words by ordering \texttt{dictionary} by the size of its centroid’s entries. A larger entry for a word implies it was more common in articles from that topic.

\texttt{julia> dictionary[sortperm(centroids[7],rev=true)]}
1000-element Array{ASCIIString,1}:
  "wind"
  "ice"
  "temperature"
  ...

(a) For each of $k = 2$, $k = 5$, and $k = 10$ run \texttt{k-means} twice, and plot $J$ (vertically) versus iteration (horizontally) for the two runs on the same plot. Create your plot by passing a vector containing the value of $J$ at each iteration to MMAPlot’s \texttt{line\_plot} function. Comment briefly on your results.

(b) Choose a value of $k$ from part (a) and investigate your results by looking at the words and article titles associated with each centroid. Feel free to visit Wikipedia if an article’s content is unclear from its title. Give a short description of the topics your clustering discovered along with the 3 most common words from each topic. If the topics do not make sense pick another value of $k$.

9. \textit{Dynamics of a compartmental system}. A \textit{compartmental system} is a model used to describe the movement of some material over time among a set of $n$ compartments of a system, and the outside world. It is widely used in \textit{pharmaco-kinetics}, the study of how the concentration of a drug varies over time in the body. In this application,
the material is a drug, and the compartments are the bloodstream, lungs, heart, liver, kidneys, and so on.

We have $n$ compartments, and we let the $n$-vector $x_t$, for $t = 1, 2, \ldots$, denote the vector of concentrations or amounts of the material in the compartments in time period $t$, so $(x_t)_i$ is the amount of the material in compartment $i$ in time period $t$. These amounts are typically, but not always, nonnegative.

We have a set of directed edges with positive weights that connect (different) compartments, or the compartments and the outside world. We let $w_{ij}$ denote the weight of the edge going from compartment $j$ to compartment $i$. If there is no edge from compartment $j$ to compartment $i$, we let $w_{ij} = 0$. We can also have edges that go from compartments to the outside world. We express these (nonnegative) edge weights as $w_{0j}$. (In the pharmaco-kinetics application, these correspond to the drug being eliminated or absorbed or otherwise rendered inactive.)

Material flows from one compartment to another, or to the outside world, over each time period, along the edges. Between time period $t$ and $t+1$, an amount of material $w_{ij}(x_t)_j$ moves along the edge from compartment $j$ to compartment $i$. This reduces the amount of material in compartment $j$ and increases it by the same amount at compartment $i$. An amount of $w_{0j}(x_t)_j$ of the material is simply removed from node $j$ between time $t$ and $t+1$. A standard assumption is that for each compartment, the sum of the weights on outgoing edges is no more than one. (What would it mean if this were not true?)

Compartmental systems are special cases of linear dynamical systems; that is, we have $x_{t+1} = Ax_t$ for some $n \times n$ matrix $A$.

A compartmental system with $n = 3$ compartments is shown below. It is a simple pharmaco-kinetic model, with compartment 1 being the bloodstream, and the periods representing 15 minute intervals (say).

(a) For the specific compartmental system shown above, find the matrix $A$ for which $x_{t+1} = Ax_t$. Be sure to account for all movement of the drug, out of and into each compartment, and elimination of the drug from the body.

(b) Assume that $x_1 = e_1$, which means we give a unit dose of the drug intravenously at period $t = 1$. Use Julia to plot the amounts of the drug in each of
the components, on the same graph, for \( t = 1, \ldots, 100 \). You can use the function `plot_compartments` in `compartmental_system.jl` to create your plot. The function accepts a \( 3 \times 100 \) matrix, where the \( i \)th row of the matrix is the amount of drug in compartment \( i \) for \( t = 1, \ldots, 100 \).

(c) Use Julia to plot the total amount of the drug in the body (the 3 compartments) versus \( t \), for \( t = 1, \ldots, 100 \). How many periods does it take for the total amount of the drug in the body to drop below \( 1/3 \) of its initial value? You can use the function `plot_totals` in `compartmental_system.jl` to create your plot. The function accepts a 100-vector that contains the total amount of drug in the body for \( t = 1, \ldots, 100 \).