DATA 1010
Problem Set 1
Due 14 September 2018 at 11 PM

Problem 1

Consider a table with 3 columns and 1000 rows, some of whose entries are missing. Denote by A the set of rows with an entry in the first column, B the set of rows with an entry in the second column, and C the set of rows with an entry in the third column. Use set notation (intersections, unions, and complements) to represent the following sets in terms of A, B, and C.

(i) The set of rows with no missing entries

(ii) The set of rows with all missing entries

(iii) The set of rows with at least one entry present

(iv) The set of rows with an entry in the first column and exactly one other entry

Problem 2

Implement the matrix multiplication algorithm from scratch in Julia (that is, any multiplication operations used must be multiplications of two *numbers*). Check your function using the following line:

```
julia> myprod([2 3 4; -4 2 5],[1 2 -4; -6 5 2; 0 1 0 ])
2×3 Array{Int64,2}:
-16  23  -2
-16   7  20
```

Problem 3

Suppose that U and V are vector subspaces of $\mathbb{R}^n$. Show that $U \cup V$ is *not* a subspace of $\mathbb{R}^n$ unless $U \subset V$ or $V \subset U$.

Problem 4

Show that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent and $\{\mathbf{v}_1 + \mathbf{w}, \mathbf{v}_2 \ldots, \mathbf{v}_n\}$ is linearly dependent, then $\mathbf{w}$ is in the span of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Problem 5

If A is a full-rank $m \times n$ matrix, then the vector in the span of the columns of A which is closest to $\mathbf{b} \in \mathbb{R}^m$ is $A\widehat{\mathbf{x}}$, where $\widehat{\mathbf{x}} = (A'A)^{-1}A'\mathbf{b}$.

Find the vector in the span of the columns of A which is closest to $\mathbf{b} = [4, 2, -1]$, where

$$A = \begin{bmatrix} 1 & -2 & 4 \\ -3 & -1 & 2 \\ 0 & 1 & -2 \end{bmatrix}$$

Problem 6

Suppose that $A = U\Sigma V'$ where Σ is diagonal and U and V are orthogonal matrices. Show that the columns of U are eigenvectors of AA' and that the columns of V are eigenvectors of $A'A$.

Hint: substitute $A = U\Sigma V'$ into the expressions AA' and $A'A$.

Problem 7

The singular value decomposition can be used to identify the primary axes in a ellipsoidal point cloud. Run the following block to generate and plot a set of 100 points.

```
using LinearAlgebra
using Plots
numpoints = 100
T = [1 2; 0 1]
P = (T * randn(2,numpoints))'
scatter(P[:,1],P[:,2],aspect_ratio=:equal)
```

Note that the coordinates of the points are stored in the rows of P.

Use Julia to compute the singular value decomposition $U\Sigma V'$ of P, and show visually that the columns of V run along the axes of the ellipse that fits the point cloud (the one shown in the figure).

Hint: `plot!([(a,b),(c,d)])` adds a line segment from the point (a,b) to the point (c,d) to the current plot. You'll want to plot line segments representing both of the columns of V.