Assignment 0 Concept Review

Linear Algebra
Gradients
Probability
Big-O Notation

Julia Overview

A0 Code Walkthrough
Linear Algebra

- Matrices $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ are denoted by upper-case letters.
- $A$ above is a $2$ by $3$ matrix ($n_{row}$ by $n_{col}$).
- Vectors $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ are denoted by lower-case letters.
- Vectors are column vectors by default ($d$ by $1$).
- Difference between $A$ and $A^T$.
- Matrix Multiplication: Computing $Ax$. 


Gradients

- Define:
  - $\mathbb{R}^d, \nabla f(x)$

- Difference between $\nabla f(x)$ and $\frac{\partial f(x)}{\partial x_j}$

- Sanity check:
  - Check the dimensions of gradient vector and input $x$
  - $f(x)$ is a scalar
  - $\nabla f(x)$ is the same dimension as $x$

- Exercise: Find the gradient
  - $f(x) = a^T x$
  - $f(x) = \log(a^T x)$
  - $f(x) = (\exp(a^T x) - 1)^3$
Probability Rules

Conditional Probability.

\[ P(A|B) = \frac{P(A \cap B)}{P(B)} \]

Bayes’ Rule.

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]

Independence.

\[ A \perp B \Rightarrow P(A|B) = P(A) \]

Marginalization.

\[ P(A) = P(A, B) + P(A, \bar{B}) \]
Rolling two dice, $D_1$ and $D_2$.

- What is $P(D_1 == 2)$?
- What is $P(D_1 + D_2 \leq 5)$?
- What is $P(D_1 == 2 \cap D_1 + D_2 \leq 5)$?
- What is $P(D_1 == 2 | D_1 + D_2 \leq 5)$?
What is $P(D_1 = 2)$?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
What is $P(D_1 + D_2 \leq 5)$?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
What is $P(D_1 = 2 \cap D_1 + D_2 \leq 5)$?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
What is $P(D_1 = 2 | D_1 + D_2 \leq 5)$?
Big-O Notation

The notation

\[ g(n) = O(f(n)) \]

means "for all large \( n \), \( g(n) \leq cf(n) \) for some constant \( c > 0 \)". Examples:

- \( 20n + 5 = O(n) \)
- \( n^2 + 50n + 10000 = O(n^2) \)
- \( 1/n + 10 = O(1) \)
- \( \log(n) + n = O(n) \)
- \( n \log(n) + n = O(n \log(n)) \)
Julia Overview

Declaring matrices, vectors, arrays

- $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is a 3x3 matrix
- $b = [1 \ 2 \ 3]$ is a row vector
- $c = [1; \ 2; \ 3]$ is a column vector

Multiplication is overloaded

- $A \times 2$ matrix-scalar
- $A \times c$ matrix-vector
- $A \times A$ matrix-matrix

Element-wise operations

- $A \times A$ matrix-matrix
- $A \times A$ element-wise multiplication
Julia Overview

Transpose

- $A'$ gives the transpose
- $c \times c$ throws error
- $c' \times c$ works

Solving linear systems

- $A \backslash b$ solves $Ax = b$
Julia Overview

Accessing elements: (use square brackets!)

- `c[1]` accesses first element of `c` (Julia is 1-indexed)
- `A[1, 2]` is scalar
- `A[2, :]` is row vector
- `A[2:3, :]` is 2-rows
- `A[2:end, :]` also works
- `A[ [1, 3], :]` non-continuous slice

Booleans

- `A .== 2` for element-wise equals
- `A .> 2` for element-wise boolean
- See: `any()`, `all()`, `find()` when boolean indexing
Julia Overview

Things of note:

▶ Use `include()` to import functions
▶ Use `readdlm()` to read files
▶ Julia passes variables by reference!
  ▶ Be careful:
    ▶ \[ x = y; \]
    ▶ \[ y[2] = 5; \]
    ▶ \[ x[2] \] is changed!
Let’s walk through the A0 code.