CS221 Section 1

Foundations

Roadmap

Python

Matrix Calculus

Recurrence Relation

Probability Theory

Syntactic Sugar

- List comprehension
- List slicing
- Passing functions
- Reading and writing files

Gotchas

- Integer division
- Tied objects
- Global variables

References

- Official Documentation (has a tutorial):

  https://docs.python.org/2.7/

- Learn X in Y minutes:

  http://learnxinyminutes.com/docs/python/

- You don't need to know `numpy`. But if you want to:

  http://nbviewer.ipython.org/gist/rpmuller/5920182

Roadmap

Python

Matrix Calculus

Recurrence Relation

Probability Theory

Useful Properties

$$"\mathbf{v} - squared" = \|\mathbf{v}\|_2^2 = \mathbf{v} \cdot \mathbf{v} = \mathbf{v}^T \mathbf{v}$$

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$$

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T$$

Matrix Calculus

$$f(\mathbf{w}) = (\mathbf{a} \cdot \mathbf{w} + 1)^2$$

Compute $\nabla_{\mathbf{w}} f(\mathbf{w})$

Matrix Calculus

$$f(\mathbf{w}) = (\mathbf{a} \cdot \mathbf{w} + 1)^2$$

Compute $\nabla_{\mathbf{w}} f(\mathbf{w})$

$$\boxed{\nabla_{\mathbf{w}} (\mathbf{a} \cdot \mathbf{w} + 1)^2 = \mathbf{2}(\mathbf{a} \cdot \mathbf{w} + \mathbf{1})\mathbf{a}}$$

A Useful Quantity

$$\boxed{\nabla_{\mathbf{w}} \mathbf{w}^\top C \mathbf{w} = (C + C^\top)\mathbf{w}}$$

Matrix Calculus

$$f(\mathbf{w}) = \|\mathbf{w}\|_2^2$$

Compute $\nabla_{\mathbf{w}} f(\mathbf{w})$

Matrix Calculus

$$f(\mathbf{w}) = \|\mathbf{w}\|_2^2$$

Compute $\nabla_{\mathbf{w}} f(\mathbf{w})$

$$\boxed{\nabla_{\mathbf{w}} \|\mathbf{w}\|_2^2 = 2\mathbf{w}}$$

Roadmap

Python

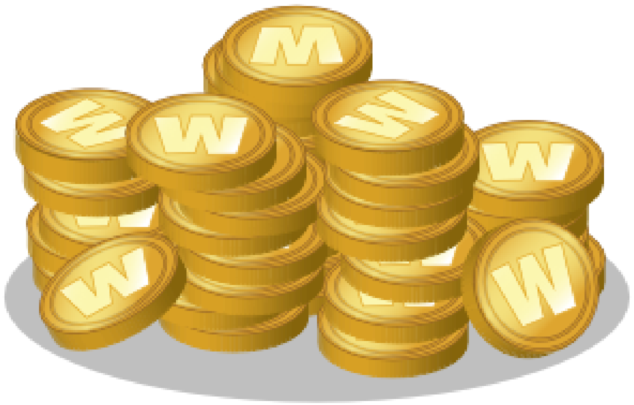
Matrix Calculus

Recurrence Relation

Probability Theory

Coin Payment

Problem



Suppose you have an unlimited supply of coins with values 2 and 3 cents

How many ways can you pay for an item costing 8 cents?

Coin Payment

Recurrence Relation: Break down into smaller problems

Memoization: Remember what you already calculated

- Refer to the extra section handout for more information regarding how the code computing this would look like.

Roadmap

Python

Matrix Calculus

Recurrence Relation

Probability Theory

Probability

Probability of event A:

$$\mathbb{P}(A)$$

Independence:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Conditional probability:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Probability

Law of total probability:

$$\mathbb{P}(A) = \sum_n \mathbb{P}(A \cap B_n) = \sum_n \mathbb{P}(B_n|A)\mathbb{P}(B_n)$$

Bayes' rule:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Random Variables

Discrete:

$$\mathbb{P}(A = a) \qquad \text{or} \qquad p_A(a)$$

Continuous:

$$\cancel{\mathbb{P}(A = a)}$$

$$f_A(a)$$

$$F(c) = \mathbb{P}(A \leq c) = \int_{-\infty}^{c} f_A(a)\, da$$

Random Variables

	A = 0	**A = 1**	**A = 2**	**A = 3**
B = 0	0.1	0.25	0.1	0.05
B = 1	0.15	0	0.15	0.2

- What is $\mathbb{P}(A = 2)$
- What is $\mathbb{P}(A = 2 \mid B = 1)$

- $\mathbb{P}(A = 2) = 0.1 + 0.15 = 0.25$
- $\mathbb{P}(A = a|B = b) = \frac{\mathbb{P}(A=a, B=b)}{\mathbb{P}(B)}$
- $\mathbb{P}(A = 2|B = 1) = \frac{0.15}{0.15+0+0.15+0.2} = 0.3$

Random Variables

Independence:

$$\forall a, b,\ \ \mathbb{P}(A = a, B = b) = \mathbb{P}(A = a)\mathbb{P}(B = b)$$

$$\forall a, b,\ \ f_{A,B}(a, b) = f_A(a) f_B(b)$$

Expectation:

$$\mathbb{E}[A] = \sum_a a\, \mathbb{P}(A = a)$$

$$\mathbb{E}[A] = \int a f_A(a) da$$

Random Variables

	A = 0	**A = 1**	**A = 2**	**A = 3**
B = 0	0.1	0.25	0.1	0.05
B = 1	0.15	0	0.15	0.2

- Are A and B independent?
- What are $\mathbb{E}[A]$, $\mathbb{E}[B]$, $\mathbb{E}[A+B]$

Linearity of Expectation: $\mathbb{E}[A+B] = \mathbb{E}[A] + \mathbb{E}[B]$

Regardless of whether A and B are independent!

- **A** and **B** are not independent. For proof, consider $\mathbb{P}(A = 0, B = 0)$, $\mathbb{P}(A = 0)$ and $\mathbb{P}(B = 0)$
- $\mathbb{E}[A] = 1.5$
- $\mathbb{E}[B] = 0.5$
- $\mathbb{E}[A + B] = 2$

Hat Toss

Problem

Suppose n hatted people toss their hats into the air and pick up one hat at random

In expectation, how many people get their own hats back?

Hint: linearity of expectation

- $X = X_1 + X_2 + ... + X_n$
- $X_i = \begin{cases} 1 & \text{if i selects own hat} \\ 0 & \text{otherwise} \end{cases}$
- $\mathbb{P}(X_i = 1) = \frac{1}{n}$
- $\mathbb{E}[X_i] = \frac{1}{n}$
- X_i are not independent, why?
- $\mathbb{E}[X] = n\frac{1}{n} = 1$

Random Variables

Variance:

$$\mathrm{Var}[A] = \mathbb{E}[(A - \mathbb{E}[A])^2] = \mathbb{E}[A^2] - \mathbb{E}[A]^2$$

Covariance:

$$\begin{aligned}\mathrm{Cov}[A, B] &= \mathbb{E}[(A - \mathbb{E}[A])(B - \mathbb{E}[B])] \\ &= \mathbb{E}[AB] - \mathbb{E}[A]\mathbb{E}[B]\end{aligned}$$

If $\mathrm{Cov}[A, B] = 0$, we say A and B are **uncorrelated**

Random Variables

If A and B are independent, then

- $\mathrm{Cov}[A, B] = \mathbb{E}[AB] - \mathbb{E}[A]\mathbb{E}[B] = 0$

  Independence implies uncorrelatedness

But the converse is **not** true!

Questions?