More Math Background

Abhishek and Matt

Sep 9, 2020

(Recall) Notation

Probability!

$$p(x) = P(X = x)$$

The name of the random variable

One of the elements in the sample space for X

"The probability that the random variable 'X' takes the value 'x'"

(Recall) Notation

Probability!

$$p(x) = P(X = x)$$

The name of the random variable

One of the elements in the sample space for X

"The probability that the random variable 'X' takes the value 'x'"

Sample Space: $\mathscr{X}$ e.g., $\mathscr{X} = \{1,2,3,4,5,6\}$

(Recall) Notation

Probability!

$$p(x) = P(X = x)$$

The name of the random variable

One of the elements in the sample space for X

"The probability that the random variable 'X' takes the value 'x'"

Sample Space: $\mathscr{X}$ e.g., $\mathscr{X} = \{1,2,3,4,5,6\}$

$\mathcal{X}$

(Recall) Notation

Probability!

$$p(x) = P(X = x)$$

The name of the random variable

One of the elements in the sample space for X

"The probability that the random variable 'X' takes the value 'x'"

Sample Space: $\mathscr{X}$ e.g., $\mathscr{X} = \{1,2,3,4,5,6\}$

(Recall) Notation

Probability!

$$p(x) = P(X = x)$$

The name of the random variable

One of the elements in the sample space for X

"The probability that the random variable 'X' takes the value 'x'"

Sample Space: $\mathscr{X}$ e.g., $\mathscr{X} = \{1,2,3,4,5,6\}$

$$\sum_{x \in \mathscr{X}} p(x) = 1$$

(Recall) Conditional Probabilities

- X = outcome of treatment
- Y = age of the patient
- P(X = Good | Y = 22) is a different value than P(X = Good | Y = 78)

$$P(X = x \mid Y = y)$$

"The probability that the random variable 'X' takes the value 'x', given that the random variable 'Y' has taken the value 'y'"

The name of another random variable

One of the elements in the sample space for Y

Example

Example

- p(eClass is down) = 0.1

Example

- p(eClass is down) = 0.1
- p(eClass is down | first day of classes) = 0.6

Example

- p(eClass is down) = 0.1
- p(eClass is down | first day of classes) = 0.6
- p(eClass is up | first day of classes) = ?

Example

- p(eClass is down) = 0.1
- p(eClass is down | first day of classes) = 0.6
- p(eClass is up | first day of classes) = ?

$$\sum_{x \in \mathcal{X}} p(x \mid Y = y) = 1$$

Sample Average

- The average of some data computed using *samples* from that data distribution

$$x_1, x_2, x_3, \ldots, x_N$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Example

Suppose you're at a mall with a friend. They're shopping, you aren't. To pass time, you decide to compute the average price of items that you see while following your friend (clearly your phone is dead).

The sequence of prices that you see are:

99, 99, 100, 99, 104, 99

What is the (sample) average?

Example

Suppose you're at a mall with a friend. They're shopping, you aren't. To pass time, you decide to compute the average price of items that you see while following your friend (clearly your phone is dead).

The sequence of prices that you see are:

99, 99, 100, 99, 104, 99

What is the (sample) average?

$$\frac{99 + 99 + 100 + 99 + 104 + 99}{6} = \frac{600}{6} = 100$$

More on sample averages

$$\frac{99 + 99 + 100 + 99 + 104 + 99}{6} = \frac{1}{6} * 99 + \frac{1}{6} * 99 + \frac{1}{6} * 100 + \frac{1}{6} * 99 + \frac{1}{6} * 104 + \frac{1}{6} * 99$$

More on sample averages

$$\frac{99+99+100+99+104+99}{6} = \frac{1}{6}*99 + \frac{1}{6}*99 + \frac{1}{6}*100 + \frac{1}{6}*99 + \frac{1}{6}*104 + \frac{1}{6}*99$$

$$= \frac{4}{6}*99 + \frac{1}{6}*100 + \frac{1}{6}*104$$

More on sample averages

$$\frac{99+99+100+99+104+99}{6} = \frac{1}{6}*99+\frac{1}{6}*99+\frac{1}{6}*100+\frac{1}{6}*99+\frac{1}{6}*104+\frac{1}{6}*99$$
$$= \frac{4}{6}*99+\frac{1}{6}*100+\frac{1}{6}*104$$
$$\approx P(X=99)*99+P(X=100)*100+P(X=104)*104$$

More on sample averages

$$\begin{aligned}\frac{99+99+100+99+104+99}{6} &= \frac{1}{6}*99+\frac{1}{6}*99+\frac{1}{6}*100+\frac{1}{6}*99+\frac{1}{6}*104+\frac{1}{6}*99 \\ &= \frac{4}{6}*99+\frac{1}{6}*100+\frac{1}{6}*104 \\ &\approx P(X=99)*99+P(X=100)*100+P(X=104)*104\end{aligned}$$

Assume all the prices in the store are among $\mathscr{X} = \{99,100,104\}$.

More on sample averages

$$\frac{99 + 99 + 100 + 99 + 104 + 99}{6} = \frac{1}{6} * 99 + \frac{1}{6} * 99 + \frac{1}{6} * 100 + \frac{1}{6} * 99 + \frac{1}{6} * 104 + \frac{1}{6} * 99$$
$$= \frac{4}{6} * 99 + \frac{1}{6} * 100 + \frac{1}{6} * 104$$
$$\approx P(X = 99) * 99 + P(X = 100) * 100 + P(X = 104) * 104$$

Assume all the prices in the store are among $\mathscr{X} = \{99,100,104\}$.

$$\text{Average(X)} = \bar{x} = \sum_{x \in \mathscr{X}} P(X = x)\ x$$

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\, x$$

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\; x$$

"What do you ***expect*** *the value of this random variable to be on average?"*

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\, x$$

*"What do you **expect** the value of this random variable to be on average?"*

- Variance

$$var(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$
$$= \sum_{x \in \mathcal{X}} p(x)(x - \mathbb{E}[X])^2$$

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\ x$$

*"What do you **expect** the value of this random variable to be on average?"*

- Variance

$$\begin{aligned} var(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \sum_{x \in \mathcal{X}} p(x)(x - \mathbb{E}[X])^2 \end{aligned}$$

"Average (squared) distance of the data points from their average"

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\, x$$

*"What do you **expect** the value of this random variable to be on average?"*

- Variance

$$var(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$
$$= \sum_{x \in \mathcal{X}} p(x)(x - \mathbb{E}[X])^2$$

"Average (squared) distance of the data points from their average"

$$\frac{1}{N} \sum_{x \in \mathcal{X}} (x_i - \bar{x})^2$$

Expectation and Variance

- Average or Expectation of a random variable X

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} p(x)\, x$$

*“What do you **expect** the value of this random variable to be on average?”*

- Variance

$$var(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_{x \in \mathcal{X}} p(x)(x - \mathbb{E}[X])^2$$

“Average (squared) distance of the data points from their average”

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\, x$$

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\, x$$

Example

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\, x$$

Example

- What is the expected value of a cup of coffee?

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\ x$$

Example

- What is the expected value of a cup of coffee?

$$\mathbb{E}[X] = \$3$$

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\, x$$

Example

- What is the expected value of a cup of coffee? $\mathbb{E}[X] = \$3$

- What is the expected value of a cup of coffee, given that it is from Starbucks?

Conditional Expectation

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \mathcal{X}} p(x \mid y)\, x$$

Example

- What is the expected value of a cup of coffee? $\mathbb{E}[X] = \$3$

- What is the expected value of a cup of coffee, given that it is from Starbucks? $\mathbb{E}[X \mid Y = \textbf{Starbucks}] = \$6$

Law of Total Expectations

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}[X \mid Y]\right]$$

$$= \sum_{y \in \mathcal{Y}} P(Y = y)\mathbb{E}[X \mid Y = y]$$

Example

- The University buys disinfecting wipes in bulk from two companies: A and B, with probability 0.3 and 0.7 each. Company A's product lasts for 50 days, Company B's, 60. What is the expected period of time a shipment will last?

Example

- The University buys disinfecting wipes in bulk from two companies: A and B, with probability 0.3 and 0.7 each. Company A's product lasts for 50 days, Company B's, 60. What is the expected period of time a shipment will last?

$$\begin{aligned}\mathbb{E}\left[D\right] &= \mathbb{E}\left[\mathbb{E}[D \mid C]\right] \\ &= P(C = A)\ \mathbb{E}[D \mid C = A] + P(C = B)\ \mathbb{E}[D \mid C = B] \\ &= 0.3 * 50 + 0.7 * 60 \\ &= 57 \text{ days}\end{aligned}$$

Vectors

"Vector"

$$\mathbf{x} \in \mathbb{R}^n$$

Vectors

(A particular)
Vector Space

$$\mathbf{x} \in \mathbb{R}^n$$

Vectors

$$\mathbf{x} \in \mathbb{R}^n$$

Examples:

$$n = 2$$

$$\mathbf{x} = \begin{bmatrix} 0.3 \\ 103.2 \end{bmatrix}$$

$$n = 3$$

$$\mathbf{y} = \begin{bmatrix} -0.2 \\ 102.023 \\ 42 \end{bmatrix}$$

Vectors

$$\mathbf{x} \in \mathbb{R}^n$$

Examples:

$$n = 2$$

$$\mathbf{x} = \begin{bmatrix} 0.3 \\ 103.2 \end{bmatrix}$$

$$n = 3$$

$$\mathbf{y} = \begin{bmatrix} -0.2 \\ 102.023 \\ 42 \end{bmatrix}$$

Indexing into a vector:

$$\mathbf{x}_1 = 0.3$$

$$\mathbf{y}_3 = 42$$

Inner (Dot) Product

$$\mathbf{x} \cdot \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{y}_i$$

Example

Simple n-dimensional example (n=2)

$$\mathbf{x} = \begin{bmatrix} -3 \\ 97 \end{bmatrix}$$

What is $\mathbf{x} \cdot \mathbf{y}$?

$$\mathbf{y} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Example

Simple n-dimensional example (n=2)

$$\mathbf{x} = \begin{bmatrix} -3 \\ 97 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

What is $\mathbf{x} \cdot \mathbf{y}$?

$$\mathbf{x} \cdot \mathbf{y} = (-3) * (-1) + (97) * (1)$$

$$= 3 + 97$$

$$\mathbf{x} \cdot \mathbf{y} = 100 = \mathbf{x}^T \mathbf{y} = < \mathbf{x}, \mathbf{y} >$$

Exercise

We have a vector space $\mathbb{R}^n$ and two vectors $\mathbf{1}, \boldsymbol{\gamma} \in \mathbb{R}^n$.
The first is a vector of ones.

$$\mathbf{1}^\top \doteq [1,1,1,\ldots].$$

The second is constructed by $\boldsymbol{\gamma}_t = \gamma^{t-1}$ where $\gamma \in [0,1)$. This results in a vector:

$$\boldsymbol{\gamma}^\top \doteq [\gamma^0, \gamma^1, \gamma^2, \ldots, \gamma^{n-1}]$$

Exercise

We have a vector space $\mathbb{R}^n$ and two vectors $\mathbf{1}, \boldsymbol{\gamma} \in \mathbb{R}^n$.
The first is a vector of ones.

$$\mathbf{1}^\top \doteq [1,1,1,\ldots].$$

The second is constructed by $\boldsymbol{\gamma}_t = \gamma^{t-1}$ where $\gamma \in [0,1)$. This results in a vector:

$$\boldsymbol{\gamma}^\top \doteq [\gamma^0, \gamma^1, \gamma^2, \ldots, \gamma^{n-1}]$$

i) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=2?

Exercise

We have a vector space $\mathbb{R}^n$ and two vectors $\mathbf{1}, \boldsymbol{\gamma} \in \mathbb{R}^n$.
The first is a vector of ones.

$$\mathbf{1}^\top \doteq [1,1,1,\ldots].$$

The second is constructed by $\boldsymbol{\gamma}_t = \gamma^{t-1}$ where $\gamma \in [0,1)$. This results in a vector:

$$\boldsymbol{\gamma}^\top \doteq [\gamma^0, \gamma^1, \gamma^2, \ldots, \gamma^{n-1}]$$

i) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=2? $1 + \gamma$

Exercise

We have a vector space $\mathbb{R}^n$ and two vectors $\mathbf{1}, \boldsymbol{\gamma} \in \mathbb{R}^n$.
The first is a vector of ones.

$$\mathbf{1}^\top \doteq [1,1,1,\ldots].$$

The second is constructed by $\boldsymbol{\gamma}_t = \gamma^{t-1}$ where $\gamma \in [0,1)$. This results in a vector:

$$\boldsymbol{\gamma}^\top \doteq [\gamma^0, \gamma^1, \gamma^2, \ldots, \gamma^{n-1}]$$

i) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=2? $1 + \gamma$

ii) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=5?

Exercise

We have a vector space $\mathbb{R}^n$ and two vectors $\mathbf{1}, \boldsymbol{\gamma} \in \mathbb{R}^n$.
The first is a vector of ones.

$$\mathbf{1}^\top \doteq [1,1,1,\ldots].$$

The second is constructed by $\boldsymbol{\gamma}_t = \gamma^{t-1}$ where $\gamma \in [0,1)$. This results in a vector:

$$\boldsymbol{\gamma}^\top \doteq [\gamma^0, \gamma^1, \gamma^2, \ldots, \gamma^{n-1}]$$

i) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=2? $1 + \gamma$

ii) What is the dot product $\langle \boldsymbol{\gamma}, \mathbf{1} \rangle$ if n=5? $1 + \gamma + \gamma^2 + \gamma^3 + \gamma^4$

Exercise

Exercise

iii) What about for any n?

Exercise

iii) What about for any n?

$$s = 1 + \gamma + \gamma^2 + \ldots + \gamma^{n-2} + \gamma^{n-1}$$

$$\gamma s = \gamma + \gamma^2 + \ldots + \gamma^{n-1} + \gamma^n$$

$$s - \gamma s = 1 - \gamma^n$$

$$s(1 - \gamma) = 1 - \gamma^n$$

$$s = \sum_{i=0}^{n-1} \gamma^i = \frac{1 - \gamma^n}{1 - \gamma}$$

Exercise

iii) What about for any n?

$$s = 1 + \gamma + \gamma^2 + \ldots + \gamma^{n-2} + \gamma^{n-1}$$

$$\gamma s = \gamma + \gamma^2 + \ldots + \gamma^{n-1} + \gamma^n$$

$$s - \gamma s = 1 - \gamma^n$$

$$s(1 - \gamma) = 1 - \gamma^n$$

$$s = \sum_{i=0}^{n-1} \gamma^i = \frac{1 - \gamma^n}{1 - \gamma}$$

iv) What happens as $\lim n \to \infty$ and $|\gamma| < 1$ (for the series to be convergent)?

Exercise

iii) What about for any n?

$$s = 1 + \gamma + \gamma^2 + \ldots + \gamma^{n-2} + \gamma^{n-1}$$

$$\gamma s = \gamma + \gamma^2 + \ldots + \gamma^{n-1} + \gamma^n$$

$$s - \gamma s = 1 - \gamma^n$$

$$s(1-\gamma) = 1 - \gamma^n$$

$$s = \sum_{i=0}^{n-1} \gamma^i = \frac{1-\gamma^n}{1-\gamma}$$

iv) What happens as $\lim n \to \infty$ and $|\gamma| < 1$ (for the series to be convergent)?

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} \gamma^i = \lim_{n\to\infty} \frac{1-\gamma^n}{1-\gamma} = \frac{1}{1-\gamma}$$

Exercise

Sample average of an incoming 'stream' of data

$$\bar{x}_N = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1} \sum_{i=1}^{N+1} x_i$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

Exercise

Sample average of an incoming ‘stream’ of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\left(\sum_{i=1}^{N} x_i + x_{N+1}\right)$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

Exercise

Sample average of an incoming 'stream' of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}(\sum_{i=1}^{N} x_i + x_{N+1})$$

$$= \frac{N}{N+1}\frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N+1}x_{N+1}$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

Exercise

Sample average of an incoming 'stream' of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}(\sum_{i=1}^{N} x_i + x_{N+1})$$

$$= \frac{N}{N+1}\frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N+1}x_{N+1}$$

$$= \frac{N}{N+1}\bar{x}_N + \frac{1}{N+1}x_{N+1}$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

Exercise

Sample average of an incoming ‘stream’ of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

$$\begin{aligned}
\bar{x}_{N+1} &= \frac{1}{N+1}\left(\sum_{i=1}^{N} x_i + x_{N+1}\right) \\
&= \frac{N}{N+1}\frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N+1}x_{N+1} \\
&= \frac{N}{N+1}\bar{x}_N + \frac{1}{N+1}x_{N+1} \\
&= \left(1 - \frac{1}{N+1}\right)\bar{x}_N + \frac{1}{N+1}x_{N+1}
\end{aligned}$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

Exercise

Sample average of an incoming ‘stream’ of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

$$\begin{aligned}
\bar{x}_{N+1} &= \frac{1}{N+1}(\sum_{i=1}^{N} x_i + x_{N+1}) \\
&= \frac{N}{N+1}\frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N+1}x_{N+1} \\
&= \frac{N}{N+1}\bar{x}_N + \frac{1}{N+1}x_{N+1} \\
&= (1 - \frac{1}{N+1})\,\bar{x}_N + \frac{1}{N+1}x_{N+1} \\
&= \bar{x}_N + \frac{1}{N+1}(x_{N+1} - \bar{x}_N)
\end{aligned}$$

Exercise

Sample average of an incoming 'stream' of data

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{x}_{N+1} = \frac{1}{N+1}\sum_{i=1}^{N+1} x_i$$

Can you compute $\bar{x}_{N+1}$ from $\bar{x}_N$?

$$\begin{aligned}
\bar{x}_{N+1} &= \frac{1}{N+1}\left(\sum_{i=1}^{N} x_i + x_{N+1}\right) \\
&= \frac{N}{N+1}\frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N+1}x_{N+1} \\
&= \frac{N}{N+1}\bar{x}_N + \frac{1}{N+1}x_{N+1} \\
&= \left(1 - \frac{1}{N+1}\right)\bar{x}_N + \frac{1}{N+1}x_{N+1} \\
&= \bar{x}_N + \frac{1}{N+1}(x_{N+1} - \bar{x}_N)
\end{aligned}$$

$$\bar{x}_{N+1} = \bar{x}_N + \alpha(x_{N+1} - \bar{x}_N)$$

Exercise

Show that $\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}[X \mid Y]\right]$

Exercise

Show that $\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}[X \mid Y]\right]$

$$= \sum_y p(y)\mathbb{E}[X \mid Y = y]$$

$$= \sum_y p(y) \sum_x p(x \mid y)\ x$$

$$= \sum_y \sum_x p(y)p(x \mid y)\ x$$

$$= \sum_x \sum_y p(y)p(x \mid y)\ x$$

$$= \sum_x x \sum_y p(y)p(x \mid y)$$

$$= \sum_x x \sum_y p(x, y)$$

$$= \sum_x x\ p(x)$$

Other Useful Maths and Resources

- Calculus:
  - https://open.umn.edu/opentextbooks/textbooks/multivariable-calculus
  - https://www.amazon.ca/Multivariable-Calculus-James-Stewart/dp/1305266641
- Probability/Statistics:
  - https://www.amazon.ca/All-Statistics-Concise-Statistical-Inference/dp/1441923225
  - http://www.utstat.toronto.edu/mikevans/jeffrosenthal/book.pdf
- Linear Algebra:
  - https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
  - https://math.mit.edu/~gs/linearalgebra/
- Machine Learning:
  - https://www.deeplearningbook.org
  - https://www.amazon.ca/Pattern-Recognition-Machine-Learning-Christopher/dp/0387310738