CNS 187 Prerequisites

- Linear Algebra
- Probability and Statistics
- Calculus (first order ODEs)
- Matlab

- Don’t worry if you’re lacking any of the above! We can guide you to books etc.
This Lecture

- Basic Probability Overview
- Basic Statistics Overview
- The Gaussian Distribution
- Parameter estimation
  - Maximum Likelihood Estimation
  - Bayesian Estimation
- Relation between Gaussian and PCA
- Signal Detection Theory and Operating Characteristics

This presentation will be put on the course website!
Probability
Example: apples and oranges

Pick a box at random (pick red box with probability 0.4), and then a fruit at random. If you picked an orange, what is the probability that you picked it from the red box?
Statistics
Expectation and Variance

\[ E(X) = \sum_{x_i} p(X = x_i) x_i \]

\[ \text{Var}(X) = \sum_{x_i} p(X = x_i) (x_i - E(X))^2 \]

\[ E(aX + bY) = aE(X) + bE(Y) \]

\[ \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \]

\[ \text{Var}(aX + b) = a^2 \text{Var}(X) \]

\[ \text{Var}(X) = E(X^2) - E(X)^2 \]
The Gaussian and PCA

\[ \mathcal{N}(x|\mu, \Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\} \]

\( x \) is a \( D \times 1 \) column vector \( x = (x_1, x_2, \ldots, x_D)^T \)

\( \mu \) is the \( D \times 1 \) mean (column) vector

\( \Sigma \) is the \( D \times D \) covariance matrix

\[ \Sigma_{i,j} = \text{Cov}(x_i, x_j) = E ((x_i - \mu_i)(x_j - \mu_j)) \]
PCA

First principal component is the eigenvector of

\[
S = \frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x})(x_n - \bar{x})^T
\]

with the largest corresponding eigenvalue \( \lambda \)

i.e. longest axis of the ellipse that defines the covariance matrix
ROC Curve

TP = true positive
FP = false positive
FN = false negative
TN = true negative

http://en.wikipedia.org/wiki/Receiver_operating_characteristic
Resources and References

- Find this presentation on the course website:

- If you want references on specific topics, please ask after class (or send an email)!

- See the last page of this presentation for a list of resources (we will reserve the books in the library)
References

- Basic Probability and Statistics

- Parameter Estimation (Maximum Likelihood, Bayesian Approach, Sufficient Statistics):

- The Gaussian Probability Distribution
  - Bishop, C.M. *Pattern Recognition and Machine Learning*. 2006. pp 78-102

- Signal Detection Theory, ROC Curves
Thank you!
Basic Rules of Probability

sum rule \[ p(X) = \sum_Y p(X, Y) \]

product rule \[ p(X, Y) = p(Y|X)p(X) \]

Bayes’ Theorem \[ p(Y|X) = \frac{p(X, Y)}{p(X)} = \frac{p(X|Y)p(Y)}{p(X)} \]