Probability in EECS

Jean Walrand – EECS – UC Berkeley
Probability in EECS - Fall 2013

Topics:
PageRank, Multiplexing, Digital Links, Speech Recognition, Path Planning, Tracking, Complements.

This course explains important applications of probability in EE and CS. The emphasis is on acquiring a working knowledge of the methodology. Homework assignments consist of Matlab experiments and problems.

Course Organization:
Lectures, one hour of discussion, two midterms (20% each), final (35%), hw (25%). We use self-grading of HW.

Prerequisites:
CS70, Math 54

Texts:
Bertsekas and Tsitsiklis: Introduction to Probability Theory;
Lecture notes provided.
Probability in EECS - Fall 2013

PageRank: How to rank search results

Google searches for "pagerank algorithm"

About 24,300,000 results (0.12 seconds)

**PageRank - Wikipedia, the free encyclopedia**

Jump to Distributed Algorithm for PageRank Computation - [edit source | edit]. There are simple and fast random walk-based distributed...

Google Panda - Google Toolbar - Webgrach - Rajeev Motwani

**Google PageRank - Algorithm**

The original PageRank algorithm was described by Lawrence Page and Sergey Brin in...

**Feature Column from the AMS**

Googlet's PageRank algorithm assesses the importance of web pages without human evaluation of the content. In fact, Google feels that the value of its service is...

**The Anatomy of a Search Engine - The Stanford University InfoLab**

PageRank or PR(A) can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web.
PageRank: Markov Chains
Multiplexing

Figure 2.1: Shared coaxial cable for internet access.

Figure 2.11: A switch with multiple input and output ports

Figure 2.16: A number of smartphones share a WiFi access point.

Figure 2.17: Norman Abramson, b. 1932.
Digital Link

Low speed, high noise  DSL, Cable Modems, etc.
Tracking

Figure 4.2: Estimating the location of a device from satellite signals.

Figure 4.14: Rudolf Kalman, b. 1930.
Speech Recognition
Path Planning

Figure 6.1: Road network. How to select a path?

Figure 6.2: A simple graph.

Figure 5.4: Richard Bellman, 1920-1984.
Complements

\[ \Gamma(i, j) = \frac{Q(i, j)}{q(i)}, j \neq i \]

Jump independent of past trajectory

Holding time independent of past trajectory

Figure 3.4: The binary symmetric channel.

Figure 7.10: Claude Shannon, 1916-2001.
## Probability in EECS - Fall 2013 - Notes

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Assignments</th>
<th>Notes</th>
<th>Book</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/29</td>
<td>PageRank</td>
<td>H1</td>
<td>1.1-3, A.1-2</td>
<td>1, 2</td>
</tr>
<tr>
<td>9/3-5</td>
<td>PageRank</td>
<td>H2</td>
<td>1.4-5</td>
<td>7.1-2, 7.</td>
</tr>
<tr>
<td>9/10-12</td>
<td>PageRank</td>
<td>H3</td>
<td>1.6</td>
<td>5.1-3, 5.</td>
</tr>
<tr>
<td>9/17-19</td>
<td>Multiplexing</td>
<td>H4</td>
<td>2.1-4, A.3</td>
<td>5.4</td>
</tr>
<tr>
<td>9/24-26</td>
<td>Multiplexing</td>
<td></td>
<td>2.5-6</td>
<td>7.3</td>
</tr>
<tr>
<td>10/1-3</td>
<td>M1 – Link</td>
<td>H5</td>
<td>3.1-2</td>
<td>8.1-2</td>
</tr>
<tr>
<td>10/8-10</td>
<td>Link</td>
<td>H6</td>
<td>3.3-6</td>
<td>9.3</td>
</tr>
<tr>
<td>10/15-17</td>
<td>Tracking</td>
<td>H7</td>
<td>4.1-3</td>
<td>8.3-4, 9.1-2</td>
</tr>
<tr>
<td>10/22-24</td>
<td>Tracking</td>
<td>H8</td>
<td>4.4-7</td>
<td></td>
</tr>
<tr>
<td>10/29-31</td>
<td>Speech</td>
<td></td>
<td>5.1-3</td>
<td></td>
</tr>
<tr>
<td>11/5-7</td>
<td>M2 – Speech</td>
<td>H9</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>11/12-14</td>
<td>Planning</td>
<td>H10</td>
<td>6.1-4</td>
<td></td>
</tr>
<tr>
<td>11/19-21</td>
<td>Planning</td>
<td>H11</td>
<td>6.5-6</td>
<td></td>
</tr>
<tr>
<td>11/26-28</td>
<td>Complements</td>
<td>H12</td>
<td>7.1-3</td>
<td></td>
</tr>
<tr>
<td>12/3-5</td>
<td>Complements</td>
<td>H13</td>
<td>7.4-6</td>
<td></td>
</tr>
<tr>
<td>12/*</td>
<td>Final</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Introduction to Probability

SECOND EDITION

Dimitri P. Bertsekas and John N. Tsitsiklis

Massachusetts Institute of Technology
Probability is common sense reduced to calculation

Laplace
Probability in EECS - Fall 2013 - Book

1. Sample Space and Probability ........................................ p. 1
   1.1. Sets .................................................................... p. 3
   1.2. Probabilistic Models ............................................ p. 6
   1.3. Conditional Probability ........................................ p. 18
   1.4. Total Probability Theorem and Bayes’ Rule ............ p. 28
   1.5. Independence ..................................................... p. 34
   1.6. Counting ........................................................... p. 44
   1.7. Summary and Discussion ...................................... p. 51
       Problems ............................................................. p. 53

2. Discrete Random Variables ............................................ p. 71
   2.1. Basic Concepts ................................................... p. 72
   2.2. Probability Mass Functions .................................. p. 74
   2.3. Functions of Random Variables .......................... p. 80
   2.4. Expectation, Mean, and Variance ......................... p. 81
   2.5. Joint PMFs of Multiple Random Variables .......... p. 92
   2.6. Conditioning ..................................................... p. 97
   2.7. Independence .................................................... p. 109
   2.8. Summary and Discussion ..................................... p. 115
       Problems ............................................................. p. 119

3. General Random Variables ........................................... p. 139
   3.1. Continuous Random Variables and PDFs .............. p. 140
   3.2. Cumulative Distribution Functions ....................... p. 148
   3.3. Normal Random Variables ................................... p. 153
   3.4. Joint PDFs of Multiple Random Variables ............ p. 158
   3.5. Conditioning ..................................................... p. 164
   3.6. The Continuous Bayes’ Rule ............................... p. 178
   3.7. Summary and Discussion ..................................... p. 182
       Problems ............................................................. p. 184
4. Further Topics on Random Variables .......................... p. 201
  4.1. Derived Distributions ........................................ p. 202
  4.2. Covariance and Correlation ................................. p. 217
  4.3. Conditional Expectation and Variance Revisited ........ p. 222
  4.4. Transforms .................................................. p. 229
  4.5. Sum of a Random Number of Independent Random Variables .... p. 240
  4.6. Summary and Discussion ................................... p. 244
      Problems ....................................................... p. 246

5. Limit Theorems ................................................... p. 263
  5.1. Markov and Chebyshev Inequalities ....................... p. 265
  5.2. The Weak Law of Large Numbers ............................ p. 269
  5.3. Convergence in Probability ................................ p. 271
  5.4. The Central Limit Theorem ................................ p. 273
  5.5. The Strong Law of Large Numbers ........................ p. 280
  5.6. Summary and Discussion ................................... p. 282
      Problems ....................................................... p. 284

6. The Bernoulli and Poisson Processes .......................... p. 295
  6.1. The Bernoulli Process ...................................... p. 297
  6.2. The Poisson Process ........................................ p. 309
  6.3. Summary and Discussion ................................... p. 324
      Problems ....................................................... p. 326
   7.1. Discrete-Time Markov Chains ................. p. 340
   7.2. Classification of States .................... p. 346
   7.3. Steady-State Behavior ....................... p. 352
   7.4. Absorption Probabilities and Expected Time to Absorption ...... p. 362
   7.5. Continuous-Time Markov Chains ............. p. 369
   7.6. Summary and Discussion .................... p. 378
       Problems ........................................ p. 380

   8.1. Bayesian Inference and the Posterior Distribution ........ p. 412
   8.2. Point Estimation, Hypothesis Testing, and the MAP Rule ...... p. 420
   8.3. Bayesian Least Mean Squares Estimation .......... p. 430
   8.4. Bayesian Linear Least Mean Squares Estimation ........ p. 437
   8.5. Summary and Discussion .................... p. 444
       Problems ........................................ p. 446

   9.2. Linear Regression .......................... p. 477
   9.3. Binary Hypothesis Testing ................. p. 486
   9.4. Significance Testing ....................... p. 496
   9.5. Summary and Discussion .................... p. 505
       Problems ........................................ p. 507

Index ................................................ p. 521
# Probability in EECS - Fall 2013 - Notes

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PageRank: Markov Chains</strong></td>
<td></td>
</tr>
<tr>
<td>1.1 Overview</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Model</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Markov Chain</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Analysis</td>
<td>12</td>
</tr>
<tr>
<td>1.5 Hitting Time</td>
<td>14</td>
</tr>
<tr>
<td>1.6 Law of Large Numbers</td>
<td>17</td>
</tr>
<tr>
<td>1.7 Summary</td>
<td>21</td>
</tr>
<tr>
<td>1.8 Problems</td>
<td>22</td>
</tr>
<tr>
<td>1.9 References</td>
<td>24</td>
</tr>
<tr>
<td><strong>Digital Link: Detection</strong></td>
<td></td>
</tr>
<tr>
<td>3.1 Overview</td>
<td>39</td>
</tr>
<tr>
<td>3.2 Detection &amp; Bayes’ Rule</td>
<td>39</td>
</tr>
<tr>
<td>3.3 Gaussian Channel</td>
<td>41</td>
</tr>
<tr>
<td>3.4 Multidimensional Gaussian Channel</td>
<td>43</td>
</tr>
<tr>
<td>3.5 Jointly Gaussian Random Variables</td>
<td>46</td>
</tr>
<tr>
<td>3.6 Hypothesis Testing</td>
<td>49</td>
</tr>
<tr>
<td>3.7 Summary</td>
<td>57</td>
</tr>
<tr>
<td>3.8 Problems</td>
<td>57</td>
</tr>
<tr>
<td>3.9 References</td>
<td>58</td>
</tr>
<tr>
<td><strong>Multiplexing: Confidence Intervals &amp; Queueing</strong></td>
<td></td>
</tr>
<tr>
<td>2.1 Overview</td>
<td>25</td>
</tr>
<tr>
<td>2.2 Sharing Links</td>
<td>26</td>
</tr>
<tr>
<td>2.3 Confidence Intervals</td>
<td>27</td>
</tr>
<tr>
<td>2.4 Central Limit Theorem</td>
<td>28</td>
</tr>
<tr>
<td>2.5 Buffers</td>
<td>33</td>
</tr>
<tr>
<td>2.6 Multiple Access</td>
<td>35</td>
</tr>
<tr>
<td>2.7 Summary</td>
<td>36</td>
</tr>
<tr>
<td>2.8 Problems</td>
<td>37</td>
</tr>
<tr>
<td>2.9 References</td>
<td>38</td>
</tr>
<tr>
<td><strong>Tracking: Estimation, Kalman Filter</strong></td>
<td></td>
</tr>
<tr>
<td>4.1 Overview</td>
<td>59</td>
</tr>
<tr>
<td>4.2 Linear Least Squares Estimates</td>
<td>60</td>
</tr>
<tr>
<td>4.3 MMSE</td>
<td>64</td>
</tr>
<tr>
<td>4.4 Vector Case</td>
<td>69</td>
</tr>
<tr>
<td>4.5 Updating LLSE</td>
<td>71</td>
</tr>
<tr>
<td>4.6 Kalman Filter</td>
<td>72</td>
</tr>
<tr>
<td>4.7 Extended Kalman Filter</td>
<td>81</td>
</tr>
<tr>
<td>4.8 Summary</td>
<td>82</td>
</tr>
<tr>
<td>4.9 Problems</td>
<td>83</td>
</tr>
<tr>
<td>4.10 References</td>
<td>84</td>
</tr>
</tbody>
</table>
A Basic Probability  
A.1 General Framework .................................. 125  
A.2 Discrete Random Variable .......................... 127  
A.3 General Random Variables ....................... 131  
A.4 Multiple Random Variables ...................... 135  
A.5 Random Vectors .......................... 140  
A.6 Problems ........................................ 141  
A.7 References ........................................ 143  

B Classification of Finite Markov Chains  
B.1 Solidarity ........................................ 145  
B.2 Invariant Distribution .............................. 146  
B.3 Periodicity ........................................ 147  
B.4 Coupling and Convergence in Distribution .... 148  

C Matlab  
C.1 Matrices ........................................ 151  
C.2 Random Variables ................................ 154  
C.3 Markov Chains .................................... 155  
C.4 Kalman Filter ..................................... 156  

D Bibliography ........................................ 159  

E Index ........................................ 161
Contents

- PageRank
- Multiplexing
- Digital Link
- Tracking
- Speech Recognition
- Path Planning
- Complements
PageRank – Search Example

How to rank search results?

---

Google Search for pagerank algorithm

About 24,300,000 results (0.12 seconds)

**PageRank - Wikipedia, the free encyclopedia**
en.wikipedia.org/wiki/PageRank
Jump to Distributed Algorithm for PageRank Computation - [edit source | edit]. There are simple and fast random walk-based distributed ...

Google Panda - Google Toolbar - Webgraph - Rajeev Motwani

**Google PageRank - Algorithm**
pr.efactory.de/e-pagerank-algorithm.shtml
The original PageRank algorithm was described by Lawrence Page and Sergey Brin in ...
Within the PageRank algorithm, the PageRank of a page T is always ...

**Feature Column from the AMS**
www.ams.org/samplings/feature-column/fcarc-pagerank
by D Austin - Related articles
Google's PageRank algorithm assesses the importance of web pages without human evaluation of the content. In fact, Google feels that the value of its service is ...

**The Anatomy of a Search Engine - The Stanford University InfoLab**
infolab.stanford.edu/~backrub/google.html
PageRank or PR(A) can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web.
Method for node ranking in a linked database

Abstract

A method assigns importance ranks to nodes in a linked database, such as any database of documents containing citations, the world wide web or any other hypermedia database. The rank assigned to a document is calculated from the ranks of documents citing it. In addition, the rank of a document is calculated from a constant representing the probability that a browser through the database will randomly jump to the document. The method is particularly useful in enhancing the performance of search engine results for hypermedia databases, such as the world wide web, whose documents have a large variation in quality.

Inventors: Page; Lawrence (Stanford, CA)
Assignee: The Board of Trustees of the Leland Stanford Junior University (Stanford, CA)
Family ID: 26673538
Appl. No.: 09/004,827
Filed: January 9, 1998
PageRank – Example

\[\pi(A) = \pi(C) + \pi(D)(1/3)\]
\[\pi(B) = \pi(A)(1/2) + \pi(D)(1/3) + \pi(E)(1/2)\]
\[\pi(C) = \pi(B) + \pi(E)(1/2)\]
\[\pi(D) = \pi(A)(1/2)\]
\[\pi(E) = \pi(D)(1/3).\]

\[\pi = [\pi(A), \pi(B), \pi(C), \pi(D), \pi(E)] = \frac{1}{39}[12, 9, 10, 6, 2].\]
PageRank – Markov Chain

\[ P[X(n+1) = j | X(n) = i, X(m), m < n] = P(i,j), \forall i, j \in \mathcal{X}, n \geq 0. \]
Let $X = k$ with probability $P(k)$, $k = 1, \ldots, K$.

How do we simulate $X$?

Example: $P(1) = 0.1$, $P(2) = 0.2$, $P(3) = 0.3$, $P(4) = 0.4$. 

$R = \text{rand}(1)$

```
Pnorm = [0, 0.1, 0.2, 0.3, 0.4]
Pcum = \text{cumsum}(Pnorm) = [0, 0.1, 0.3, 0.6, 1]
R = \text{rand}(1) \rightarrow 0.52
[a, X] = \text{histc}(R, Pcum) \rightarrow a = [0, 0, 1, 0, 0], X = 3
```
Let $X = k$ with probability $P(k)$, $k = 1, \ldots, K$.

How do we simulate $X$?

Example: $P(1) = 0.1$, $P(2) = 0.2$, $P(3) = 0.3$, $P(4) = 0.4$.

```matlab
function T = discrete(P)
    % This function generates a
    % random variable equal to k w.p. P(k)/sum(P).
    % Here, P = [P(1), P(2), ..., P(K)] where
    % the P(k) are nonnegative.
    Pnorm=[0 P]/sum(P);
    Pcum=cumsum(Pnorm);
    R=rand(1);
    [~,T] = histc(R,Pcum);
```
M = 100;
A = 1;
P = [0, 0.3, 0.7; 0, 0.4, 0.6; 1, 0, 0];
X = zeros(1,M);
X(1) = A;
for m = 1:M-1
    X(m+1) = discrete(P(X(m,:),:));
end
function simMC(M,A,P)
% simulate a MC for M steps with initial state A and tpm P
% 
X = zeros(1,M);
X(1) = A;
for m = 1:M-1
    X(m+1) = discrete(P(X(m),:));
end
plot(X,'.-b')
PageRank – Markov Chain - Simulation

```
>> M = 100; A = I; P = [0.9, 0.1; 0.1, 0.9];
>> simMC(M,A,P)

>> P = [0.7, 0.3; 0.3, 0.7];
>> simMC(M,A,P)
```
PageRank – Markov Chain - Simulation

\[
P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
\]

\[
\text{simMC(M,A,P)}
\]

\[
P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}
\]

\[
\text{simMC(M,A,P)}
\]
>> P=[0,0.3,0.7;0, 0.4,0.6;1,0,0];
>> simMC(M,A,P)
PageRank – Markov Chain - Simulation

\[ P=\begin{bmatrix} 0,0.5,0,0.5,0;0,0,1,0,0;1,0,0,0,0;1/3,1/3,0,0,1/3;0,0.5,0.5,0,0 \end{bmatrix}; \]

\[ \text{>> simMC(M,A,P)} \]
PageRank – Markov Chain – Inv. Dist.

Invariant Distribution

\[ \pi = \pi P \leftarrow \text{Balance Equations} \]

\[
P = \begin{bmatrix}
1 - a & a \\
b & 1 - b
\end{bmatrix}
\]

\[ \pi(P - I) = 0 \]

\[
\pi \begin{bmatrix}
-a & a \\
b & -b
\end{bmatrix} = [0, 0]
\]

\[
\pi \begin{bmatrix}
-a & 1 \\
b & 1
\end{bmatrix} = [0, 1]
\]

\[ \pi = [0, 1] \begin{bmatrix}
-a & 1 \\
b & 1
\end{bmatrix}^{-1} = \left[ \frac{b}{a + b}, \frac{a}{a + b} \right] \]
Invariant Distribution

\[ \pi = \pi P \leftrightarrow \text{Balance Equations} \]

\[ \pi(P - I) = [0, \ldots, 0]; \]
\[ \pi B = [0, \ldots, 0, 1] \text{ where } B = [(P - I)_1, \ldots, (P - I)_{n-1}, 1] \]
\[ \pi = [0, \ldots, 0, 1]B^{-1} \]
PageRank – Markov Chain – Inv. Dist.

\[ \pi = [0.3077 \quad 0.2308 \quad 0.2564 \quad 0.1538 \quad 0.0513] \]