Course Info

- [http://courses.soe.ucsc.edu/courses/tim250/Spring13/01](http://courses.soe.ucsc.edu/courses/tim250/Spring13/01)

- **Schedule**
  - Mid term and end of term
Performance Evaluation

Final Exam (closed book):
Week 11 of the Quarter

Performance Evaluation:
Homework and class participation 30%
Midterm 20% (Week 6 of the Quarter)
iPinYouProject 30%
Final Exam 20% (Week 11 of the Quarter)
Audience Participation
Questionnaire

• **Background**
  – Industry/Academia
  – Major
  – Programming experience

• **Expectations from taking TIM250**
Brief Bio James G. Shanahan

- 25 years in the fields of data mining and operations research
  - Principal and Founder, Boutique Data Consultancy
    - Clients include: Adobe, Digg, Adobe, SearchMe, AT&T, SkyGrid, TapJoy
  - Affiliated with University of California Santa Cruz (UCSC, ISM250,251,209)
  - Chief Scientist, Turn Inc. (A CPX ad network, DSP)
  - Principal Scientist, Clairvoyance Corp (CMU spinoff; sister lab to JRC)
  - Co-founder of Document Souls (an anticipatory information system)
  - Research Scientist, Xerox Research
  - Research Engineer, Mitsubishi Group
  - PhD in machine learning (1998), University of Bristol, UK;
    B.Sc. Comp. Science (1989), Uni. of Limerick, Ireland

- Areas of expertise include:
  - Local search, web search, online advertising, machine learning, ranking,
    user modeling, statistics, social networks
Disclaimer

• The Authors retains all rights, including copyrights and distribution rights.
• No publication or further distribution in full or in part permitted without explicit written permission from the author
• Living vicariously!
TIM250, Spring 2013, Section 01: Home

TIM 250: Stochastic Optimization in Information Systems and Technology with a special focus on Digital Online Advertising

TIM 250 in brief

<table>
<thead>
<tr>
<th>TIM 250</th>
<th>Stochastic Optimization in Information Systems and Technology with a special focus on Digital Online Advertising</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dates</td>
<td>April 2, 2013 – May, 2013</td>
</tr>
<tr>
<td>Time</td>
<td>Tuesdays 6:00 pm - 9:30 pm</td>
</tr>
<tr>
<td>Location</td>
<td>Main UCSC campus JB156 and Silicon Valley Center (SVC) SVC 303</td>
</tr>
<tr>
<td>Webpage</td>
<td><a href="http://www.soe.ucsc.edu/courses/course?ism250">http://www.soe.ucsc.edu/courses/course?ism250</a> <a href="http://www.soe.ucsc.edu/courses/course/tim250/Spring13/01">http://www.soe.ucsc.edu/courses/course/tim250/Spring13/01</a></td>
</tr>
</tbody>
</table>
Course Description

• Essentially, every problem in business, computer science and engineering can be formulated as the optimization of some function under some set of constraints.
  – This universal reduction automatically suggests that such optimization tasks are intractable. Fortunately, most real world problems have special structure, such as convexity, locality, decomposability or submodularity.
  – These properties allow us to formulate optimization problems that can often be solved efficiently.
  – As such operations research is a quantitative discipline that deals with the application of advanced analytical methods to help make better decisions.
  – The terms management science and decision science are sometimes used as more modern-sounding synonyms. Employing techniques from other mathematical sciences, such as mathematical modeling, statistical analysis, and mathematical optimization, operations research arrives at optimal or near-optimal solutions to complex decision-making problems.
Applications of OR abound!

- Techniques from operations research are leveraged within many industries, in both online and offline modes, and are responsible for both batch and real-time decision-making involving trillions of dollars ($10^{12}$) annually worldwide. For example:
  - Finance: How Wall Street manages stock portfolios?
  - Digital Advertising: How publishers optimize advertiser revenue online?
  - Logistics: How FedEx ships billions of packages annually across 220 countries?
  - Social networks: How to suggest people you may know?
  - Web search: How to rank web pages with out queries?
### Some applications of operations research

<table>
<thead>
<tr>
<th>Organization</th>
<th>Area of Application</th>
<th>Section</th>
<th>Annual Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federal Express</td>
<td>Logistical planning of shipments</td>
<td>1.3</td>
<td>Not estimated</td>
</tr>
<tr>
<td>Continental Airlines</td>
<td>Reassign crews to flights when schedule disruptions occur</td>
<td>2.2</td>
<td>$40 million</td>
</tr>
<tr>
<td>Swift &amp; Company</td>
<td>Improve sales and manufacturing performance</td>
<td>3.1</td>
<td>$12 million</td>
</tr>
<tr>
<td>Memorial Sloan-Kettering Cancer Center</td>
<td>Design of radiation therapy</td>
<td>3.4</td>
<td>$459 million</td>
</tr>
<tr>
<td>United Airlines</td>
<td>Plan employee work schedules at airports and reservations offices</td>
<td>3.4</td>
<td>$6 million</td>
</tr>
<tr>
<td>Welch’s</td>
<td>Optimize use and movement of raw materials</td>
<td>3.6</td>
<td>$150,000</td>
</tr>
<tr>
<td>Samsung Electronics</td>
<td>Reduce manufacturing times and inventory levels</td>
<td>4.3</td>
<td>$200 million more revenue</td>
</tr>
<tr>
<td>Pacific Lumber Company</td>
<td>Long-term forest ecosystem management</td>
<td>6.7</td>
<td>$398 million NPV</td>
</tr>
<tr>
<td>Procter &amp; Gamble</td>
<td>Redesign the production and distribution system</td>
<td>8.1</td>
<td>$200 million</td>
</tr>
<tr>
<td>Canadian Pacific Railway</td>
<td>Plan routing of rail freight</td>
<td>9.3</td>
<td>$100 million</td>
</tr>
<tr>
<td>United Airlines</td>
<td>Reassign airplanes to flights when disruptions occur</td>
<td>9.6</td>
<td>Not estimated</td>
</tr>
<tr>
<td>U.S. Military</td>
<td>Logistical planning of Operations Desert Storm</td>
<td>10.3</td>
<td>Not estimated</td>
</tr>
<tr>
<td>Air New Zealand</td>
<td>Airline crew scheduling</td>
<td>11.2</td>
<td>$6.7 million</td>
</tr>
<tr>
<td>Taco Bell</td>
<td>Plan employee work schedules at restaurants</td>
<td>11.5</td>
<td>$13 million</td>
</tr>
<tr>
<td>Waste Management</td>
<td>Develop a route-management system for trash collection and disposal</td>
<td>11.7</td>
<td>$100 million</td>
</tr>
<tr>
<td>Bank Hapoalim Group</td>
<td>Develop a decision-support system for investment advisors</td>
<td>12.1</td>
<td>$31 million more revenue</td>
</tr>
<tr>
<td>Sears</td>
<td>Vehicle routing and scheduling for home services and deliveries</td>
<td>13.2</td>
<td>$42 million</td>
</tr>
<tr>
<td>Conoco-Phillips</td>
<td>Evaluate petroleum exploration projects</td>
<td>15.2</td>
<td>Not estimated</td>
</tr>
<tr>
<td>Workers’ Compensation Board</td>
<td>Manage high-risk disability claims and rehabilitation</td>
<td>15.3</td>
<td>$4 million</td>
</tr>
<tr>
<td>Westinghouse</td>
<td>Evaluate research-and-development projects</td>
<td>15.4</td>
<td>Not estimated</td>
</tr>
<tr>
<td>Merrill Lynch</td>
<td>Manage liquidity risk for revolving credit lines</td>
<td>16.2</td>
<td>$4 billion more liquidity</td>
</tr>
<tr>
<td>PSA Peugeot Citroën</td>
<td>Guide the design process for efficient car assembly plants</td>
<td>16.8</td>
<td>$130 million more profit</td>
</tr>
<tr>
<td>KeyCorp</td>
<td>Improve efficiency of bank teller service</td>
<td>17.6</td>
<td>$20 million</td>
</tr>
<tr>
<td>General Motors</td>
<td>Improve efficiency of production lines</td>
<td>17.9</td>
<td>$90 million</td>
</tr>
<tr>
<td>Deere &amp; Company</td>
<td>Management of inventories throughout a supply chain</td>
<td>18.5</td>
<td>$1 billion less inventory</td>
</tr>
<tr>
<td>Time Inc.</td>
<td>Management of distribution channels for magazines</td>
<td>18.7</td>
<td>$3.5 million more profit</td>
</tr>
<tr>
<td>Bank One Corporation</td>
<td>Management of credit lines and interest rates for credit cards</td>
<td>19.2</td>
<td>$75 million more profit</td>
</tr>
<tr>
<td>Merrill Lynch</td>
<td>Pricing analysis for providing financial services</td>
<td>20.2</td>
<td>$50 million more revenue</td>
</tr>
<tr>
<td>AT&amp;T</td>
<td>Design and operation of call centers</td>
<td>20.5</td>
<td>$750 million more profit</td>
</tr>
</tbody>
</table>

**Other Areas**
- Digital Advertising
- Web search
Class Topics

• More concretely, this class serves as a first graduate course in optimization with an emphasis on theory, intuition, and problem solving in management and engineering.

• This class will focus on problem formulation, software technologies and analytical methods for optimization serving as an introduction to a wide variety of optimization problems and techniques including:
  – linear programming, nonlinear programming, dynamic programming, network flows, integer programming, heuristic approaches, Markov chains, game theory, and decision analysis.
TIM 250 Focus

- TIM250 will focus on getting students familiar with core principles in Stochastic Optimization, grounding these principles in both
  - We will develop the techniques systematically and sequentially, but will move back-and-forth the different real world applications and domain areas.
  - (1) examples taken primarily form online advertising (a $65 Billion industry) and in
  - (2) example projects and code in R and Matlab. Each class will be composed of theory, practice and problems, thereby informing and inspiring students on how to apply theory to practice.
TIM 250 Background

• Prerequisites:

• Students should have some background in probability, statistics, calculus, and linear algebra, though most of the concepts from these areas, which are core to optimization, will be reviewed for completeness during the course.
TIM 250 is timely!

• **TIM 250**
  – Followon to TIM 206 (Optimization theory and applications)
  – Goes deeper into NLP, MDPs, Convex optimization, Machine learning as optimization, and digital advertising and TIM 251 (scalable, distributed decision support systems)

• **.... with applications in digital advertising and online marketing, healthcare, airlines, financial**

• **Timely:**
  – Growing flood of online data, Budding industries (e.g., digital advertising, healthcare, financial)
  – Computational power is available (PC, Cloud computing, Hadoop)
  – Progress in algorithms and theory and applications
Summaries → Decisions → Personalization

• The old days were about asking, ‘What is the biggest, smallest, and average?’” says Michael Olson, CEO of startup Cloudera. “Today it’s, ‘What do you like? Who do you know?’ It’s answering these complex questions.”

• Personalization through optimization
  – Healthcare; finance; marketing and advertising

• In the old days:
  – A retailer such as Macy’s (M) that once pored over last season’s sales information could shift to looking instantly at how an e-mail coupon for women’s shoes played out in different regions.
Data Scientists

Hacking Skills

Math, OR Probability Theory

Domain Expertise

Digital Advertising & Marketing, Econometrics, Web Search, Healthcare, Finance

Statistics, Optimization Theory, Social Network Analytics, Geoinformational Science

Hadoop, Java, Python, R
150,000 Data Scientists needed in US

[McKinsey Report on Big Data]
Data-Driven Decision Making is hot skill
Data-Driven Decision Making is a hot skill.

b hunters with these IT skills are assured of employment, now and in the future.

ry Brandel

y 11, 2007 (Computerworld) Have you spoken with a high-tech recruiter or professor of computer science lately? According to observers across the country, the technology skills shortage that pundits were ringing about a year ago is real (see “Workforce crisis: Preparing for the coming IT crunch”).

iving everything I see in Silicon Valley is completely contrary to the assumption that programmers are a dying breed and being offshore,” says Kevin Scott, senior engineering manager at Google Inc. and a founding member of the professions and education boards at the Association for Computing Machinery. “From big names to start-ups, companies are hiring as aggressively as possible.”

o check out our updated 8 Hottest Jobs for ’08.

ny recruiters say there are more open positions than they can fill, and according to one recruiter of IT at Icetake University in Milwaukee, students are getting snapped up before they graduate. In January, Kaiser asked 34 students in the systems analysis 1 design class she was teaching how many had already accepted offers to begin work after graduating in May. Twenty-four of them raised their hands. “I feel sure that 10 who didn’t have all been given offers,” she says.

vice it to say, the market is strong, but only if you have the right skills. I want to be part of the companies that hire what eight experts — curriculum developers and skills of the near future say. See also “The top 10.

Machine learning

Machine learning is the process that allows computers to learn from data and improve their accuracy. As companies work to build software such as collaborative filtering, spam filtering and fraud-detection applications that seek patterns in jumbo-size data sets, some observers are seeing a rapid increase in the need for people with machine-learning knowledge, or the ability to design and develop algorithms and techniques to improve computers’ performance, Scott says.

“It’s not just the case for Google,” he says. “There are lots of applications that have big, big, big data sizes which creates a fundamental problem of how you organize the data and present it to users.”

Done
Operations Research has even gotten hip!

• It’s not going too far to say that data analytics and optimization has even gotten hip.
  – The San Francisco offices of startup Splunk have all the of-the-moment accoutrements you’d find at Twitter or Zynga.

• The engineers work in what amounts to a giant living room with pinball machines, foosball tables, and Hello Kitty-themed cubes.

• Weekday parties often break out—during a recent visit, it was Mexican fiesta. Employees were wearing sombreros and fake moustaches while a dude near the tequila bar played the bongos.
• …Advertising and Marketing
Supply

DEMAND
Advertiser wishes to reach consumers

SUPPLY
- Sponsored Search
- Contextual
- Display
- Classified
- Email
- Social
- Mobile

Consumers
Sponsored Search, Display Volumes

DEMAND
Advertiser wishes to reach consumers

SUPPLY
- Sponsored Search
  - 208 Billion in US (2009)
- Display

Advertiser  
Agency  
Ad

Publish

Consumers
From Mad Men To Wall Street and beyond!

<table>
<thead>
<tr>
<th>Banner</th>
<th>Click+Data</th>
<th>Personal</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Set in New York City, Mad Men begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.</td>
<td>2007</td>
<td>Data</td>
<td></td>
</tr>
<tr>
<td>Human Intensive Lots of guess work Forward Market</td>
<td>Increasingly</td>
<td>Technology Data Driven Forward Market Spot Markets</td>
<td></td>
</tr>
<tr>
<td>YoY: Double digit growth</td>
<td>Advertisers still in broadcast mode</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1st Generation 2nd Generation 3rd Generation
Sponsored Search: Ad Impression

Ads are clearly distinguishable from the actual search results and they rotate.
Sponsored Search: Click

Google

data mining

Search

Advanced Search
Preferences

Web Groups Scholar Books

Personalized Results 1 - 10 of about 66,300,000 for data mining [definition]. (0.14 seconds)

Data Mining Software
FREE: 30-day Eval & Online Training Webcast, Guided Tour, Case Studies

Mine Text Data
Analyze Consumer Opinions
Categorize Issues Automatically
www.clarabridge.com

Data mining - Wikipedia, the free encyclopedia
Data mining can be defined as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data". [1] Data mining may ...
en.wikipedia.org/wiki/DataMining - 68k - Cached - Similar pages - Note this

Data Mining: What is Data Mining?
Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into ...
www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm - 13k - Cached - Similar pages - Note this

Data Mining Techniques
Data Mining is an analytic process designed to explore data (usually large amounts of data - typically business or market related) in search of consistent ...
www.statsoft.com/textbook/stdatadmin.html - 47k - Cached - Similar pages - Note this

Data Mining: Text Mining, Visualization and Social Media

Easy Data Mining
Discover a data mining system that easily exports data to Excel.
Datawatch.iresponse.net

Data Mining Software
Discover insights hidden in your existing data using SPSS solutions.
www.spss.com
Impressions \(\rightarrow\) Clicks \(\rightarrow\) Transactions

\[
\text{Impressions} \rightarrow \text{Clicks} \rightarrow \text{Transactions}
\]

\[
\begin{align*}
\text{CPC} & \iff \text{RPC} \iff \text{Revenue} \\
\text{ctr} & = \frac{\text{clicks}}{\text{impressions}} \\
\text{cost} & = \text{clicks} \times \text{cpc} \\
\text{cpa} & = \frac{\text{cost}}{\text{conversions}} \\
\text{rpc} & = \frac{\text{revenue}}{\text{clicks}} \\
\text{profit} & = \text{revenue} - \text{cost} \\
\text{margin} & = \frac{\text{profit}}{\text{cost}}
\end{align*}
\]
Forecasting

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Match Type</th>
<th>Max CPC</th>
<th>Imp</th>
<th>Clicks</th>
<th>CTR</th>
<th>Avg CPC</th>
<th>Cost</th>
<th>Avg Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>tennis shoes</td>
<td>Broad</td>
<td>$3.11</td>
<td>13456</td>
<td>234</td>
<td>1.74%</td>
<td>$2.98</td>
<td>$697.32</td>
<td>3</td>
</tr>
<tr>
<td>running shoes</td>
<td>Exact</td>
<td>$0.27</td>
<td>4356</td>
<td>26</td>
<td>0.60%</td>
<td>$0.13</td>
<td>$3.38</td>
<td>4</td>
</tr>
<tr>
<td>sneakers</td>
<td>Phrase</td>
<td>$1.17</td>
<td>2234</td>
<td>34</td>
<td>1.52%</td>
<td>$1.15</td>
<td>$39.10</td>
<td>5.8</td>
</tr>
<tr>
<td>best running shoes</td>
<td>Broad</td>
<td>$2.67</td>
<td>198755</td>
<td>345</td>
<td>0.17%</td>
<td>$2.27</td>
<td>$783.15</td>
<td>3.2</td>
</tr>
<tr>
<td>basketball shoes</td>
<td>Exact</td>
<td>$0.13</td>
<td>13</td>
<td>0</td>
<td>0.00%</td>
<td>$ -</td>
<td>$ -</td>
<td>4.2</td>
</tr>
</tbody>
</table>
## Maximize Revenue: Ad Allocation Example

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AudienceZone 1</td>
<td>Volume_{11}</td>
<td>Volume_{ij}</td>
<td>Volume_{ij}</td>
<td>Volume_{ij}</td>
<td>35</td>
</tr>
<tr>
<td>AudienceZone 2</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>50</td>
</tr>
<tr>
<td>AudienceZone 3</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>15</td>
</tr>
<tr>
<td>Demand Contracted PageViews</td>
<td>45</td>
<td>20</td>
<td>30</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Use Linear Programming to generate suggested Ad Sales schedule; This schedule shows how much volume is available in each audience zone and how it should be allocated to advertiser targeting criteria to maximize publisher revenue. E.g., This schedule shows how much of Audience zone 1 should be sold to advertisers who want to target sport people; this optimized volume, Volume_{11} (based on a global picture) will generate a corresponding Revenue_{11}
The optimal portfolios plotted along the curve have the highest expected return possible for the given amount of risk.
LP Healthcare Example: Design of External Beam Radiation Therapy

• MARY has just been diagnosed as having a cancer at a fairly advanced stage. Specifically, she has a large malignant tumor in the bladder area (a “whole bladder lesion”).

• The goal of the design is to select the combination of beams to be used, and the intensity of each one, to generate the best possible dose distribution.

• Cross section of Mary’s tumor (viewed from above)
  – as well as nearby critical tissues to avoid and
  – the radiation beams being used.
  – normally dozens of possible beams must be considered

• Location of her tumor is in a tricky spot.
Design of Radiation Therapy is Key

• Because of the need to carefully balance all these factors, the design of radiation therapy is a very delicate process.

• The goal of the design
  – is to select the combination of beams to be used, and
  – the intensity of each one, to generate the best possible dose distribution.

  (The dose strength at any point in the body is measured in units called kilorads.)

• Once the treatment design has been developed, it is administered in many installments, spread over several weeks.
F(Beam intensity) = Absorption

IsoDose Map

- The contour lines represent the dose strength as a percentage of the dose strength at the entry point. A fine grid then is placed over the isodose map. By summing the radiation absorbed in the squares containing each type of tissue, the average dose that is absorbed by the tumor, healthy anatomy, and critical tissues can be calculated. (measured in Kilorads)
F(Beam intensity) = Absorption

For any proposed beam of given intensity, the analysis of what the resulting radiation absorption by various parts of the body would be requires a complicated process.

In brief, based on careful anatomical analysis, the energy distribution within the twodimensional cross section of the tissue can be plotted on an isodose map, where the contour lines represent the dose strength as a percentage of the dose strength at the entry point. A fine grid then is placed over the isodose map. By summing the radiation absorbed in the squares containing each type of tissue, the average dose that is absorbed by the tumor, healthy anatomy, and critical tissues can be calculated. With more than one beam (administered sequentially), the radiation absorption is **additive (cross product terms)**.
Founded in 1837, Deere & Company is a leading worldwide producer of equipment for agriculture, forestry, and consumer use. The company employs approximately 43,000 people and sells its products through an international network of independently owned dealers and retailers.

For decades, the Commercial and Consumer Equipment (C&CE) Division of Deere pushed inventories to the dealers, booked the revenues, and hoped that the dealers had the right products to sell at the right time. However, the division had an inventory-to-annual-sales ratio of 58 percent based on inventories at Deere and at its dealers in 2001, so inventory costs were getting badly out of control. Ironically, although dealers had large inventories, they often did not have the right products in stock.

C&CE’s supply chain managers needed to cut inventory levels while improving product availability and delivery performance. They had read about inventory optimization successes in Fortune, so they hired a leading OR consulting firm (SmartOps) to tackle this challenge. With 300 products, 2,500 North American dealers, five plants and associated warehouses, seven European warehouses, and several retailers’ consignment warehouses, the coordination and optimization of C&CE’s supply chain was indeed a formidable challenge.

However, SmartOps rose to this challenge very successfully by applying state-of-the-art inventory optimization techniques embedded in its multistage inventory planning and optimization software product to set trustworthy targets. C&CE used these targets, together with appropriate dealer incentives, to transform the operation of its entire supply chain on an enterprise-wide basis. In the process, Deere improved its factories’ on-time shipments from 63 percent to 92 percent, while maintaining customer service levels at 90 percent. By the end of 2004, the C&CE Division also had exceeded its goal of $1 billion of inventory reduction or avoidance.

Amazon: A logistics provider
Irhythm: detect cardiac problems

• IRhythm makes a type of oversize, plastic band-aid called the Zio Patch that helps doctors detect cardiac problems before they become fatal.
• Patients affix the Zio Patch to their chests for two weeks to measure their heart activity. The patients then mail the devices back to IRhythm’s offices, where a technician feeds the information into Amazon’s cloud computing service.
• Patients typically wear rivals’ much chunkier devices for just a couple of days and remove them when they sleep or shower—which happen to be when heart abnormalities often manifest. The upside of the waterproof Zio Patch is the length of time that people wear it—but 14 days is a whole lot of data.
Sensors + Services => Privacy Problem

- Personal devices (with GPS’ and accelerometers)
  - Earphones; Nike+ (measures and records the distance and pace of a walk or run); asthma inhaler with built-in GPS tracking

- Personal/social services
  - Mint, Twitter, diets, health, exercise, FaceBook

- These data streams create a huge privacy problem
Always connected at the extreme → Lifelogging

Records events using multiple wearable sensors
Provides access to these data at multiple levels of granularity and abstraction, using access mechanism based on the episodic memory of human beings.

http://www.imrc.kist.re.kr/wiki/LifeLog
Backend Technology
3rdi Art Project

- A New York University arts, Professor Bilal
- A surgically-implanted camera (12/15/2010)
  - 3rdi Project, has already generated international media attention and anticipation. On Dec. 15 images from the "third eye" in the back of Bilal's head -- a surgically-implanted camera -- will be unveiled in Doha, Qatar as part of the Told/Untold/Retold exhibition that inaugurates the new Arab Museum of Modern Art near Education City, Doha’s intellectual hub.

- Transmits one image per minute to a website (www.3rdi.me), displayed a Doha gallery
  - with the inaugural images to be unveiled in a custom-designed room in the Doha gallery of the museum’s new permanent collection, 20 years in the making, including more than 6,000 works by Arab artists from North Africa to the Gulf, from the 1920s to the present day.
4 Screens: Mobile, Computer, TV, Theatre

- **Smartphones 50% share in mid2011 (US)**
- **Tablet computers**
  - Large Format Benefit
  - Enhanced mobile apps
  - Total media tablets device market
  - 28MM in 2011 (ABI, 2010; Barclays Capital, 2010)
- **IPTV**
  - Play IPTV digital content originating from the iTunes Store, Netflix, YouTube, Flickr, MobileMe or any Mac OS X or Windows computer running iTunes onto an enhanced-definition or high-definition widescreen television
  - Still early days but
- **Theatre**
## Personalized Marketing

### 1st Generation
- Human Intensive
- Lots of guess work
- Forward Market

### 2nd Generation
- Technology
- Data Driven
- Forward Market
- Spot Markets

### 3rd Generation
- Advertisers still in broadcast mode

### 2007
- Data
- Personalization

**YoY: Double digit growth**

*Set in New York City, *Mad Men* begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.*
Machine Learning

• **Supervised ML**
  – Constrained and unconstrained optimisation

• **Unsupervised ML**
  – Constrained and unconstrained optimisation

• **Reinforcement Learning**
  – Markov processes
The Netflix competition pitted teams from across the world to find ways to predict movies better for users.

The Netflix data set includes:
- Integral ratings from 1 to 5 for 17,000 movies
- From 480,000 users
- Totalling over 100 million ratings

Essentially a sparse $480000 \times 17000$ matrix

Total entries: roughly 8.5 billion. Matrix is close to 99% empty!
Netflix Data: Users and Videos

Idea: “Compress” each dimension of the matrix (users $\times$ movies) to approximate the known values. Each dimension is then represented by a finite number of features which are combined to estimate the missing values.
Using SVD For Matrix Approximation

If approximation or compression is desired, a lower-rank SVD $\tilde{M}$ can be computed from $M$:

- Rank $r$ is chosen and only the $r$ singular values of greatest magnitude are used.
- Thus the rank of $U$ and $V$ are reduced as well.
- $U$, $\Sigma$, and $V$ are reduced to dimension $m \times r$, $r \times r$, and $r \times n$ respectively.

This approximation minimizes the Frobenius norm of $A = M - \tilde{M}$:

$$\| A \|_F = \sqrt{\sum_{i=1}^{\min\{m, n\}} \sigma_i^2} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$$
Incremental SVD

- An incremental algorithm for SVD was developed by Simon Funk based on a paper by Genevieve Gorrell, in which the SVD is approximated numerically via gradient descent.
- An approximation rank $r$ is chosen beforehand such that there are $r$ feature vectors for movies and users.
- The $\sum$ matrix is left blended into the feature vectors, so only $U$ and $V$ remain.
- Initial values for these feature vectors are set arbitrarily, i.e. 0:1.
- Gradient descent is used to minimize the function for each feature vector along users and movies.
Partial Matrix SVD (some empty cells)

Fill in empty cells using $U$ and $V$ factorization

$\sim R_{1,3} = \langle U^T_1, V_3 \rangle$
Matrix Factorization has many Free Parameters

- Matrix Factorization in Collaborative Filtering
  - To fit the product of two (low rank) matrices to the observed rating matrix.
  - To find two latent user and item feature matrices.
  - To use the fitted matrix to predict the unobserved ratings.

\[
\begin{pmatrix}
  u_{11} & \cdots & u_{1k} \\
  \vdots & \ddots & \vdots \\
  u_{m1} & \cdots & u_{mk}
\end{pmatrix}
\begin{pmatrix}
  v_{11} & \cdots & v_{1n} \\
  \vdots & \ddots & \vdots \\
  v_{k1} & \cdots & v_{kn}
\end{pmatrix}
\]

User-specific latent feature vector

Item-specific latent feature column vector

Netflix: 500,000 users, 17,000 movies = 8.5Billion ratings; only 100Million provided
20 million free parameter for Rank-40 SVD Decomposition: \( 5 \times 10^5 \times 40 + (40 \times 17000) \)
Regularized SVD-like Decomposition

- Minimize the loss based on the observed ratings with regularization terms to avoid over-fitting problem

\[
\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_i^T V_j)^2 + \frac{\lambda_1}{2} \|U\|^2_F + \frac{\lambda_2}{2} \|V\|^2_F
\]

where \( \lambda_1, \lambda_2 > 0 \).

- The problem can be solved by simple gradient descent algorithm.

- Alternatively, instead of using a simple linear factor model (or bounded version as above), the inner product between user-specific and item-specific feature vectors is mapped through a nonlinear logistic function:
  - \( g(x) = 1/(1+\exp(-x)) \), which bounds the range of the predictions into \([0, 1]\).
Gradient

\[ E = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} I_{ij} (V_{ij} - p(U_i, M_j))^2 + \frac{k_u}{2} \sum_{i=1}^{n} \|U_i\|^2 + \frac{k_m}{2} \sum_{j=1}^{m} \|M_j\|^2, \]

When using the prediction function (3), the objective function and its negative gradients have the following forms:

\[ E = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} I_{ij} (V_{ij} - p(U_i, M_j))^2 \]
\[ + \frac{k_u}{2} \sum_{i=1}^{n} \|U_i\|^2 + \frac{k_m}{2} \sum_{j=1}^{m} \|M_j\|^2, \]

\[ -\frac{\partial E}{\partial U_i} = \sum_{j=1}^{m} I_{ij} ((V_{ij} - p(U_i, M_j))M_j) - k_u U_i, i = 1, \ldots, n, \]

\[ -\frac{\partial E}{\partial M_j} = \sum_{i=1}^{n} I_{ij} ((V_{ij} - p(U_i, M_j))U_i) - k_m M_j, j = 1, \ldots, m. \]

One can then perform the optimization of \(U\) and \(M\) by gradient descent:
Over 300-400 applications of ML on this page
More Data versus Rocket Science

Some simple math using a mountain of data can get you 80% of the way!

Figure 2. Learning Curves for Confusable Disambiguation
The Data Knows!

http://www.businessweek.com/magazine/data-analytics-crunching-the-future-09082011.html
Optimization is core to the
Big Data Revolution

- 30 billion
  Pieces of content shared on Facebook every month
- 5 billion
  Mobile phones in use in 2010
- $600 billion
  Potential annual consumer surplus from personal location data globally
- 60%
  Potential increase in retailers’ operation margins with big data
Course Modus Operandi

- Course will focus on getting students familiar with core principles in operations research.
- Grounding these principles in both
  - (1) examples taken primarily from online advertising (a $70 Billion industry), healthcare, Airline industry, etc.
  - And in (2) example projects and code in R/Excel/IO OR Solver.
- Each class will be composed of theory, geometry/code, and problems, thereby informing and inspiring students on how to apply theory to practice.
Some Practical Skills

- Problem solving
- Data analysis
- Coding up algorithms
- Real-world datasets
- Evaluations and metrics
- Collaboration
- Presentation
- Teamwork
Performance Evaluation

Final Exam (closed book):
Week 11 of the Quarter

Performance Evaluation:
Homework and class participation 30%
Midterm 20% (Week 6 of the Quarter)
iPinYouProject 30%
Final Exam 20% (Week 11 of the Quarter)
Audience Participation
Questionnaire

• **Background**
  – Industry/Academia
  – Major
  – Programming experience

• **Expectations from taking TIM206**
Course philosophy

• **Socratic Method (both inspiration and information)**
  – participation strongly encouraged (please state your name and affiliation)

• **Highly interactive and adaptable**
  – Questions welcome!!

• **Lectures emphasize intuition, rigor and detail**
  – Build on lectures
  – Background reading will provide more rigor & detail

• **Action Items**
  – Read suggested books first (and then papers), read/write Wikipedia, watch/make YouTube videos, take other courses, participate in competitions, do internships, network
  – Prototype, simulate, publish, participate
  – Classic (core) versus trendy (applications)
<table>
<thead>
<tr>
<th>Week</th>
<th>Core Subject</th>
<th>Application Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Online advertising, Ad networks, real-time bidding (RTB), IPinYou data science competition ($160,000 prize)</td>
<td>Perceptron learning; linear regression; in R/Matlab code</td>
</tr>
<tr>
<td>2</td>
<td>Gradient Descent, Lines, Tangents, Taylor Series, Slopes/Gradients, Roots of an equation, Newton-Rhapson</td>
<td>Perceptron learning; linear regression; in R/Matlab code</td>
</tr>
<tr>
<td>3</td>
<td>Classical Programming, Equality Constraints, Convexity, concavity, constraint optimization, Lagrange Multipliers</td>
<td>Media Planning in online advertising</td>
</tr>
<tr>
<td>4</td>
<td>Nonlinear Programming, first/second order conditions, efficient frontier</td>
<td>Keyword portfolio management</td>
</tr>
<tr>
<td>5</td>
<td>Kuhn Tucker Conditions, primal from, dual form</td>
<td>Support Vector Machines</td>
</tr>
<tr>
<td>6</td>
<td>Linear Programming, Standard form, augmented form, Simple algorithm, Interior point method, slack variables</td>
<td>Forecasting and ad display scheduling</td>
</tr>
<tr>
<td>7</td>
<td>Dynamic Programming; Markov Decision Processes</td>
<td>Optimal frequency capping policy for ads</td>
</tr>
<tr>
<td>8</td>
<td>Online/sequential Learning; Gittins index, Regret, pacing algorithms, proportional control</td>
<td>Multi-arm bandit approach to ad optimization</td>
</tr>
<tr>
<td>9</td>
<td>Stochastic Recommenders</td>
<td>Predicting Click thru rates</td>
</tr>
<tr>
<td>10</td>
<td>Project presentations: Students present their findings</td>
<td>IPinYou data science competition</td>
</tr>
<tr>
<td>11</td>
<td>Exam</td>
<td></td>
</tr>
</tbody>
</table>
JGS Lecture Outline

- Introduction and some motivating problems
- Linear Algebra Basics Review
- Fundamental theorem of LP
- Matrix-view and the fundamental insight
- Duality
- Interior point Algorithm
- Transportation Problem
- Applying linear programming to online advertising
- Summary
Books and References

  - This book is available for purchase from the UCSC bookstore (in eBook form at a negotiated reduced price of $80 compared with the print version prices of $250; please use ISBN 9781121774254 for purchasing) and comes with very useful software.
  - There should be copies available in the Library also.
- https://highered.mcgraw-hill.com/sites/0073376299/student_view0/
- Other refs will be available on website as PDFs
Objective Function: Max \[ Z = 5x_1 + 3x_2 \]

Constraint:
- \[ 0 \leq 0x_1 + 0x_2 \leq 0 \]
- \[ 1 \]
- \[ 0 \leq 5x_1 \leq 9 \]
- \[ 0 \leq 3x_1 + 4x_2 \leq 12 \]

Note:
The number of constraints can't be more than 5
Subject to:
- \[ x_1 \geq 0, x_2 \geq 0 \]
- \[ 5x_1 \leq 9 \]
- \[ 3x_1 + 4x_2 \leq 12 \]

Optimum

After the optimal solution is reached, you can do sensitivity analysis by pressing the button below.

Sensitivity Analysis
Online Resources: Optimization in R

- [http://cran.r-project.org/web/views/Optimization.html](http://cran.r-project.org/web/views/Optimization.html)

- **NLP SUMT**
  - R sumt(...)

- SUMT examples and good write up
  - [http://www.me.utexas.edu/~jensen/ORMM/supplements/units/nlp_methods/const_opt.pdf](http://www.me.utexas.edu/~jensen/ORMM/supplements/units/nlp_methods/const_opt.pdf)
  - The methods that we describe presently, attempt to approximate a constrained optimization problem with an unconstrained one and then apply standard search techniques to obtain solutions. The approximation is accomplished in the case of penalty methods by adding a term to the objective function that prescribes a high cost for violation of the constraints. In the case of barrier methods, a term is added that favors points in the interior of the feasible region over those near the boundary. For a problem with n variables and m constraints, both approaches work directly in the n-dimensional space of the variables. The discussion that follows emphasizes penalty methods recognizing that barrier methods embody the same principles.
Course References

- R
  - Practical Regression and Anova using R, [http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf](http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf), by JJ Faraway (please download PDF)
    - Chapter 1 - Getting Started With R (PDF available)
    - Chapter 6 - Diagnosing Problems in Linear and Generalized Linear Models
R abline() example

?abline.

## Setup up coordinate system (with x==y aspect ratio):
plot(c(-2,3), c(-1,5), type = "n", xlab="x", ylab="y", asp = 1)
## the x- and y-axis, and an integer grid
abline(h=0, v=0, col = "gray60")
text(1,0, "abline( h = 0 )", col = "gray60", adj = c(0, -.1))
abline(h = -1:5, v = -2:3, col = "lightgray", lty=3)
abline(a=1, b=2, col = 2)
text(1,3, "abline( 1, 2 )", col=2, adj=c(-.1,-.1))
R Basics

example.GettingStarted.Chapter1.Fox()

- **R via a GUI** R Commander
  - Examine data; plot data

- **Scripting in R**
  - Variables, vectors, data.frames, functions, graphics

- **Check out example.GettingStarted.Chapter1.Fox()**
R Commander

```
data()
Hist(Duncan$education, scale="frequency", breaks="Sturges", col="darkgray")
.Table <- table(Duncan$Type)
.Table # counts for type 100*.Table/sum(.Table) # percentages for type remove(.Table)
boxplot(Duncan$education, ylab="education")
  (boxplot(income-type, ylab="income", xlab="type", data=Duncan)
  ) (boxplot(prestige-type, ylab="prestige", xlab="type", data=Duncan)
scatter3d(Duncan$education, Duncan$income, Duncan$prestige, fit="linear", residuals=TRUE, bg="white", axis.scales=TRUE, grid=TRUE, ellipsoid=FALSE, 
xlab="education", ylab="income", zlab="prestige")
```
R Basics

example.GettingStarted.Chapter1.Fox()

• **R via a GUI**  R Commander  
  – Examine data; plot data

• **Scripting in R**  
  – Variables, vectors, data.frames, functions, graphics

• **Check out example.GettingStarted.Chapter1.Fox()**
Simple Plotting Example

- # Example 1
- # make a very simple plot
  
x <- c(1,3,6,9,12)
y <- c(1.5,2,7,8,15)
plot(x,y)
Plotting in R: plot character

- Plot symbols are set within the `plot()` function by setting the `pch` parameter (plot character?) equal to an integer between 1 and 25.
x <- c(1,3,6,9,12)
y <- c(1.5,2,7,8,15)

# Example 2. Draw a plot, set a bunch of parameters.
plot(x,y, xlab="x axis", ylab="y axis", main="my plot",
     ylim=c(0,20), xlim=c(0,20), pch=15, col="blue")
Plotting examples

par(mfrow=c(2,3))

plot(x, type="p", main="plot(x,type="p")") # Note the escaped quotes \\
plot(x, type="l", main="plot(x, type="l")")
plot(x, type="b"~, main="plot(x, type="b")")
plot(x, type="h"~, main="plot(x, type="h")")
plot(x, type="s", main="plot(x, type="s")")
plot(x, type="n", main="plot(x, type="n")")
Different symbols and line types

```r
par(mfrow=c(2,2))
# Different symbols and line types
plot(x, pch="x")
plot(x, type="l", lty=2)
plot(x, pch="x", cex=2)
plot(x, type="l", lwd=2)
```
x <- c(1,3,6,9,12)
y <- c(1.5,2,7,8,15)

# Example 2. Draw a plot, set a bunch of parameters.
plot(x,y, xlab="x axis", ylab="y axis", main="my plot",
     ylim=c(0,20), xlim=c(0,20), pch=15, col="blue")

# fit a line to the points
myline.fit <- lm(y ~ x)

# get information about the fit
summary(myline.fit)

# draw the fit line on the plot
abline(myline.fit)
# Example 3
# add some more points to the graph
x2 <- c(0.5, 3, 5, 8, 12)
y2 <- c(0.8, 1, 2, 4, 6)
points(x2, y2, pch=16, col="green")

my plot
• The `text()` function allows us to put text on the plot where we want it. An obvious use is to label a line or group of points.

```r
text(c(2,2),c(37,35),labels=c("Non-case","Case"))
```
# Example 1
# make a very simple plot
x <- c(1,3,6,9,12)
y <- c(1.5,2,7,8,15)
plot(x,y)
text(c(3,10),c(3,10),labels=c("Case1","Case4"),
     col="red")

The text() function allows us to put text on the plot where we want it.
Plotting Example with margin text

# Example 1
# make a very simple plot
x <- c(1,3,6,9,12)
y <- c(1.5,2,7,8,15)
plot(x,y)
text(c(3,10),c(3,10),labels=c("Case1","Case4"),
col="red")

Text labels can also be placed in the margins of a plot using the `mtext()` function. This would place the words "Low" and "High" on the second line below the X axis centered at 3 and 10 units.
Two y-axis example

# http://rgraphics.limnology.wisc.edu/line.php

rm(list = ls())  # Clear all variables
graphics.off()   # Close graphics windows

# Generate sample time series data
ti = 1:50          # Generate 50 sample time steps
# Generate 50 stochastic data points for time series y1
y1 = 8 + rnorm(50)

# Plot the y1 data
par(oma=c(2,2,2,4))   # Set outer margin areas (only necessary in order to plot extra y-axis)
plot(ti, y1,
     type="b",  # Data to plot - x, y
     main="Time series plot",  # Main title for the plot
     xlab="Time",    # Label for the x-axis
     ylab="Response (y1 & y2)",  # Label for the y-axis
     font.lab=2,     # Font to use for the axis labels: 1=plain text, 2=bold, 3=italic, 4=bold italic
     ylim=c(0,20),   # Range for the y-axis; "xlim" does same for x-axis
     xaxp=c(0,50,5),  # X-axis min, max and number of intervals; "yaxp" does same for y-axis
     bty="l")       # Box around plot to contain only left and lower lines

http://rgraphics.limnology.wisc.edu/line.php
# Add y2 data to the same plot
points(ti, y2, 
  type="b", # Plot lines and points
  lty=1,   # Line type: 0=blank, 1=solid, 2=dashed, 3=dotted, 4=dotdash, 5=longdash, 6=twodash
  lwd=1,  # Line width
  pch=19, # Point type: pch=19 - solid circle, pch=20 - bullet (smaller circle), pch=21 - circle, pch=22 - square, pch=23 - diamond, pch=24 - triangle point-up, pch=25 - triangle point down.
  col="red") # Color of the plotted data

# Add y3 data to the same plot, but on a different axis
par(new=T, # The next high-level plotting command (actually plot.new) should not clean the frame before drawing ias if it was on a new device.
  oma=c(2,2,2,4)) # Increase the size of the outer margins to accommodate second y axis

plot(ti, y3, 
  yaxt="n", # Do not plot the y-axis
  ylab="", # Do not plot the y-axis label
  xlab="", # Do not plot the x-axis label
  type="b", # Plot lines and points
  lty=1,   # Line type: 0=blank, 1=solid, 2=dashed, 3=dotted, 4=dotdash, 5=longdash, 6=twodash
  lwd=1,  # Line width
  pch=19, # Point type: pch=19 - solid circle, pch=20 - bullet (smaller circle), pch=21 - circle, pch=22 - square, pch=23 - diamond, pch=24 - triangle point-up, pch=25 - triangle point down.
  col="blue") # Color of the plotted data

axis(4, # Add a second axis: 1=below, 2=left, 3=above and 4=right
  pretty(range(y3),10)) # Intervals for the second y-axis
mtext("Response (y3)", # Add second y-axis label
  side=4, # Add to right hand side of plot
  line=3, # Add to line 3 from the margin
  font=2) # Print label in bold

# Add a legend to the plot
legend("topleft", # x-y coordinates for location of the legend
  legend=c("y1", "y2", "y3"), # Legend labels
  col=c("black", "red", "blue"), # Color of points or lines
  pch=c(21,19,19), # Point type
  lty=c(1,1,1), # Line type
  lwd=c(1,1,1), # Line width
  legend="Time series") # Legend title
Two y-axis example

http://rgraphics.limnology.wisc.edu/line.php
fx=function(x) {
  2*x^3 - 3*x^2 - 12*x + 6
}

x=seq(-4, 4, by=0.1)
plot(x, fx(x), main="my graph", xlab="x",
    ylab="2*x^3 - 3*x^2 - 12*x + 6", pch="", " type="l")
grid()
R Graphics Basics: Save plots to PDF

### R code and examples for "Modern Applied Statistics Using R"
### Lecture 3: Graphics
### Alexander.Ploner@ki.se  2007-09-17

# The basic high-level plot
x = rnorm(25)
y = 2 + 3*x + rnorm(25)
plot(x)

# Create a pdf file in home directory
setwd("~")
pdf("test.pdf")
plot(x)
dev.off()
# Nice trick - works if a pdf viewer is installed
viewer = options()$pdfviewer
system(paste(viewer, "test.pdf"))

### Note: we can easily create multipage plots
pdf("test2.pdf")
plot(x, main="Page 1")
plot(y, main="Page 2")
dev.off()

[http://www.meb.ki.se/~aleplo/R2007/Rcourse03.R] See example.PDF()
Useful for reporting
Putting it all together ...(PIAT)

```r
x = rnorm(25)
y = 2 + 3*x + rnorm(25)
par(mfrow=c(1,2))
# Common text elements
plot(x, main="Changing titles and labels", s
  xlab="Index of observation", ylab="Observa

# Adding extra points and lines; we switch to
plot(x,y)
points(mean(x), mean(y), pch="X", cex=2, f
lines(range(x), range(y))

# Adding text and arrows
text(max(x), min(y), "Center", col=2, adj=c(1
arrows(max(x)-strwidth("Center"), min(y), mean(x), mean(y), col="red",
  lwd=2)
```
Changing titles and labels

Index of observation
Extra subtitles go here

Center
More Graphic Examples

#incorporate more examples from http://www.meb.ki.se/~aleplo/R2007/Rcourse03.R

###
Given a Pt. and Slope...

\[ f(x) = x^3 - 12x + 1 \]

First derivative
\[ f'(x) = 3x^2 - 12 \]

\[ f(x) = 0 \text{ at maximum and minimum} \]

Using \((x_1, f(x_1))\) and \(m = f'(x_1)\)

And the equation formula
\[ y - y_0 = m(x - x_0) \]

Plot the tangent line

Given a Pt. and Slope... Approximate \(f(x)\) with tangent
Approximate a curve using a Tangent

- Given a point on the curve, \((x_0, f(x_0))\) and a slope, \(f'(x_0)\), we can calculate the equation of the tangent at \((x_0, f(x_0))\) as follows:
  - \[ y - y_0 = f'(x_0)(x - x_0) \]  
  - \[ f(X) - f(x_0) = f'(x_0)(X-x_0) \] where \(X\) is a free variable, \(f'(x)\) is the slope
  - Then for any \(X\) in the neighbourhood of \(X_0\) we can approximate it by the tangent at \((x, f(x_0))\)
  - Of course it will not be that accurate but can be reasonably approximate if \(X\) is not too far from \(x_0\).
Exercise 1.1

• Given function \( f(x) = x^3 - 12x + 1 \) approximate the curve at \((-1, f(-1))\) in the \(x\) range of \([-3, 3]\) using the tangent to \((-1, f(-1))\) [also know as the first order Taylor approximation]

• In R, plot the curves \( f(x) \), \( f'(x) \) and the tangent approximation and label appropriately

• Add text and arrows to highlight \((-1,f(-1))\) and its tangent line

• Comment on the approximation of \( f(x) \) at \( x = -3 \)
  \[ f_{\text{Tangent}}(x = -3) \]

• HINT: review material on slides before this and after this.
Derivatives in R using deriv(), D()

\begin{verbatim}
fx=function(x) {
    2*x^3 - 3*x^2 - 12*x + 6
}
fprime = function (x){
    6*x^2 - 6 * x - 12
}
dx2x <- deriv(~ x^2, "x", TRUE)
> dx2x
function (x)
{
    .value <- x^2
    .grad <- array(0, c(length(.value), 1L), list(NULL, c("x")))
    .grad[, "x"] <- 2 * x
    attr(.value, "gradient") <- .grad
    .value
}
\end{verbatim}

\begin{itemize}
    \item \(F(x) = x^2\); \(f'(x) = 2x\); \(\frac{df(x)}{dx} = f'(x)\)
    \item See example.drawTangent()
\end{itemize}

> dx2x(2)
4
> dx2x <- deriv(~ x^2+x+1, "x", TRUE)
> attr(dx2x(2), "gradient")
Derivatives in R using deriv(), D()

F(x) = X^2; f’(x) = 2x; df(x)/dx=f’(x)

> dx2x <- deriv(~ x^2, "x", TRUE) ; dx2x

function (x)
{
  .value <- x^2
  .grad <- array(0, c(length(.value), 1L), list(NULL, c("x")))
  .grad[, "x"] <- 2 * x
  attr(.value, "gradient") <- .grad
  .value
}

See example.drawTangent()

C:\jimi\Projects\R\GradientDescent\JimisMLCourse.R
Tangent Example: \( f(x), f'(x) \)

\[
f(x) = 2x^3 - 3x^2 - 12x + 6
\]

\[
f'(x) = 6x^2 - 6x - 12
\]

Given a point on the curve and a slope:

1. Choose \( x \) at index 10
2. \( x = \text{seq}(-4, 4, \text{by}=0.1) \)
3. \( i = 11 \)
4. \( x0 = x[i] \)
5. \( f_x0 = f(x0) \) or \( y0 \)
6. \( \text{slope} = f\text{prime}(x0) \)
7. \( \text{tangentLine} = \text{function} (x, \text{slope}, x0, y0) \) \{ 
   \[
   y = \text{slope} \times (x-x0) + y0
   \]
\}
8. \( y = \text{tangentLine}(x, \text{slope}, x0, f_x0) \)
9. \( \text{par}(\text{mfrow} = c(2, 1)) \) # split display region into 2 rows and one column (i.e., 2 regions)
10. \( x = \text{seq}(-4, 4, \text{by}=0.1) \)
11. \( \text{plot}(x, f(x), \text{main}="2x^3 - 3x^2 - 12x + 6", \text{xlab}="x", \text{ylab}="f(x)"), \text{pch}="" ) \)
12. \( \text{lines}(x, f(x), \text{lt}=1, \text{col}="\text{red}" ) \)
13. \( \text{lines}(x, y, \text{col}="\text{blue}" ) \)
14. \( \text{points}(x0, f_x0, \text{col}="\text{blue}" , \text{bg}="\text{blue}" , \text{pch}=21) \)

\( f'(x) \) or \( f\text{prime}(x) \)

\( f'(x) \) or \( f\text{prime}(x) \)

Approx curve using tangent at \((x0, f(x0))\)

tangent given \((x0, f(x0))\) and slope

\( f'(x) \) or \( f\text{prime}(x) \)
Tangent Example: plotting

```r
par(mfrow = c(2, 1)) # split display region into to 2 rows and one column (i.e., 2 regions)
x=seq(-4, 4, by=0.1)
plot(x, fx(x), main="2*x^3 - 3*x^2 - 12*x + 6", xlab="x", ylab="f(x)", pch="")
lines(x, fx(x), lty=1, col="red")
lines(x, y,col="blue")
points(x0, f_x0, col="blue", bg="blue", pch=21)
text(x0-0.5, f_x0+10, paste("(", x0, ", ", f_x0, ")", sep="\n"))
grid()

plot(x, fprime(x), main="f'(x), i.e., slope of tangent to f(x) at x
f'(x)=6*x^2 - 6 * x - 12",
ylab="fprime(x)", pch="")
lines(x, fprime(x), lty=3, col="red") #hand calculate slope
lines(x, attr(fprime.deriv(x), "gradient"), lty=2, col="green") #R-calculated slope
text(x0-0.5, slope+10, paste("("", x0, ", ", slope, ")", sep="\n"))

points(x0, slope, col="blue", bg="blue", pch=21)
grid()
```

See example.drawTangent()
Exercise 1.1

The point of contact of the tangent and the curve is (-3, -39)

Slope is 60 (f'(-3)=60)
Summary on Tangent Approximations

\[ y - y_1 = m(x - x_1) \] give a point \((x_1, y_1)\) and a slope \(m\)

\[ f(x) - y_1 = f'(x_1)(x - x_1) \] let \(y = f(x)\) and \(m = f'(x_1)\)

\[ f(x) = y_1 + f'(x_1)(x - x_1) \]

\[ f(x) = f(x_1) + f'(x_1)(x - x_1) \] eqn of tangent at point \((x_1, f(x_1))\) given \((x_1, f(x_1))\) and slope = \(f'(x_1)\)

- Remember
  - Every point on the curve has a tangent
  - A tangent is a straight line
  - The tangent has its own equation
  - The tangent has equation \(y = mx + c\)
  - This equation is different for every position of the tangent since the slope \((f'(x))\) is different.
A more complicated example

\[ Y = mx + c \]

\[ y - y_0 = m(x - x_0) \]

Let \( m = f'(x_0) \)
Algorithms and Courseware

- **OR Courseware is a program called OR Tutor**
  - personal tutor to help you learn the algorithms. It consists of many demonstration examples that display and explain the algorithms in action.

- **OR Courseware includes many interactive routines for executing the algorithms interactively in a convenient spreadsheet format.**
  - The computer does all the routine calculations while you focus on learning and executing the logic of the algorithm.

- **Microsoft Excel: formulate small OR models in a spreadsheet format. The Excel Solver then is used to solve the models.**
  - Your OR Courseware includes a separate Excel file for nearly every chapter in this book. Each time a chapter presents an example that can be solved using Excel, the complete spreadsheet formulation and solution is given in that chapter’s Excel file. For many of the models in the book, an Excel template is available also.

- **R is an open source programming language and software environment for statistical computing and graphics.**

- **CPLEX is an elite state-of-the-art software package that is widely used for solving large and challenging OR problems**
On Windows 8 (via Parallels on Mac): Press Command Key for a couple of Seconds to bring up Windows VM
Then Right Click to bring up APPs
Make sure Java is installed

Free Java Download

Download Java for your desktop computer now!

Version 7 Update 10

Free Java Download

» What is Java? » Do I have Java? » Need Help?

Why download Java?

Java technology allows you to work and play in a secure computing environment.

IORTutorial (interactive) and OR Tutor
Objective Function: \[ Z = 5x_1 + 3x_2 \]

Note:
The number of constraints can’t be more than 5
Subject to:
\[ x_1 \geq 0, \ x_2 \geq 0 \]
(1) \[ 5x_1 \leq 9 \]
(2) \[ 3x_1 + 4x_2 \leq 12 \]

After the optimal solution is reached, you can do sensitivity analysis by pressing the button below.

Optimum
\[ Z = 13.95 \text{ with } x_1 = 1.8, \ x_2 = 1.65 \]
The Graphical Method - Formulation

Using money units of thousands of dollars and time units of hours, the data for Leo's problem can be summarized as follows.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Resource Usage per Share of Each Venture</th>
<th>Amount of Resource Available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Venture 1</td>
<td>Venture 2</td>
</tr>
<tr>
<td>Money</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>Time</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Savings per share</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

Since Leo must decide how many shares of each venture to take, the decision variables are

\[ \begin{align*}
\varphi_1 & = \text{number of shares of Venture 1,} \\
\varphi_2 & = \text{number of shares of Venture 2.}
\end{align*} \]

He wants to maximize his savings, so the value of the objective function is

\[ Z = \text{number of money units (thousands of dollars) added to savings}. \]

The linear programming model for Leo's problem now can be formulated from the data in the table as follows.

Maximize \[ Z = 20\varphi_1 + 10\varphi_2 \]

subject to

\[ \begin{align*}
\varphi_1 - \varphi_2 & \leq 1 \\
3\varphi_1 + \varphi_2 & \leq 7 \\
\varphi_1 & \geq 0, \quad \varphi_2 & \geq 0.
\end{align*} \]
**General Linear Programming**

1. Graphical Method
2. Slack Variables
3. Algebraic Simplex
4. Tabular Simplex
5. Fundamental Insight
6. Sensitivity Analysis

---

**The Graphical Method**

Using money units of thousands of dollars, the data for Leo's problem can be summarized as follows:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Amount of Resource Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money</td>
<td>1</td>
</tr>
<tr>
<td>Time</td>
<td>7</td>
</tr>
<tr>
<td>Savings</td>
<td>3</td>
</tr>
</tbody>
</table>

Since Leo must decide how many of each resource to use, we have:

\[
\begin{align*}
\alpha_1 &= \text{amount of resource 1} \\
\alpha_2 &= \text{amount of resource 2}
\end{align*}
\]

He wants to maximize his savings, so we have:

\[
Z = 3\alpha_1 + 7\alpha_2
\]

The linear programming model can be stated from the data in the table as follows:

subject to

---

**Network Analysis**

**Integer Programming**

**Nonlinear Programming**

**Markov Decision Processes**

**Simulation**
Phases of an OR Study

1. Define the problem of interest Gather relevant data.
2. Formulate a mathematical model to represent the problem.
3. Build model:
   1. Develop a computer-based procedure for deriving solutions to the problem from the model.
4. Debug/evaluate the model and refine it as needed.
5. Build supporting infrastructure for using such a model in production (accessing data, logging data, make decisions, etc.)
6. Deploy (and ABTest).
Systems Modeling is inherently interactive and iterative

1. Understand the domain and Define problems
2. Formulate solution/model
3. Derive solutions
4. Debug/Evaluate models
5. Build supporting infrastructure for models
6. Deploy System in the wild (and AB test)
Systems Modeling is inherently interactive and iterative

1. Understand domain, Collect requirements & Data
2. Collect requirements, and Data
3. Feature Engineering
4. Modeling: Extract Patterns/Models
5. Interpret and Evaluate discovered knowledge
6. Deploy System in the wild (and AB test)

6 Steps to Data Modeling in Practice
Step 1: Define the Problem and get data

- In contrast to textbook examples, most practical problems encountered by OR teams are initially described to them in a vague, imprecise way.
- Therefore, the first order of business is to study the relevant system and develop a well-defined statement of the problem to be considered.
- This includes determining such things as the appropriate objectives, constraints on what can be done, interrelationships between the area to be studied and other areas of the organization, possible alternative courses of action, time limits for making a decision, and so on.
- This process of problem definition is a crucial one because it greatly affects how relevant the conclusions of the study will be. It is difficult to extract a “right” answer from the “wrong” problem!

- Beer flow problem (Taps versus Point of Sale)
  - Problem
  - Data
Step 2: Formulate a mathematical model

- Formulate a mathematical model to represent the problem.
- Mathematical models are idealized representations, that are expressed in terms of mathematical symbols and expressions.
- Thus, if there are \( n \) related quantifiable decisions to be made, they are represented as decision variables (say, \( x_1, x_2, \ldots, x_n \)) whose respective values are to be determined.
- The appropriate measure of performance (e.g., profit) is then expressed as a mathematical function of these decision variables:
  \[ P = 3x_1 + 2x_2 + 5x_n \]
  This function is called the objective function.
- Restrictions on variables can be captured as constraints:
  \[ 2x_1 + x_2 + 5x_n \leq 10 \]
- Parameters are the constants:
  - The constants (namely, the coefficients and right-hand sides) in the constraints and the objective function are called the parameters of the model.
Decision Variables vs. Parameters

• Use Linear Programming as an example
  – Define problem
  – Gather data
  – Formulate model
  – Solve

Maximize $Z = 3X_1 + 5X_2$
Subject to:

Parameters

$1X_1 \leq 4$
$2X_2 \leq 12$
$3X_1 + 2X_2 \leq 18$

Decision Variables

SOLVE
COMPUTE

LP(c, a, b) $\rightarrow$ x
x: Decision variables
$c$: Objective coefficients
$a$:
b: RHS resource constraints

$X_1 = 2, X_2 = 6$
Determining the parameter values

• Determining the appropriate values to assign to the parameters of the model (one value per parameter) is both a critical and a challenging part of the model-building process.

• In contrast to textbook problems where the numbers are given to you, determining parameter values for real problems requires gathering relevant data.

• Gathering accurate data frequently is difficult. Therefore, the value assigned to a parameter often is, of necessity, only a rough estimate.

• Sensitivity analysis helps simulate many possible parameters settings
  – Because of the uncertainty about the true value of the parameter, it is important to analyze how the solution derived from the model would change (if at all) if the value assigned to the parameter were changed to other plausible values.

• E.g., Maximize profit (what if sale price fluctuates; cost of raw materials fluctuates)
Step 3: Deriving Solutions from the Model

Plug and chug

• Apply a computer-based procedure for deriving solutions to the problem from the mathematical model.

• Sometimes, in fact, it is a relatively simple step, in which one of the standard algorithms (systematic solution procedures) of OR is applied on a computer by using one of a number of readily available software packages.

10-20% of modeling effort

• For experienced OR practitioners, finding a solution is the fun part, whereas the real work comes in the preceding and following steps, including the postoptimality analysis.

• Course: How to find solutions to various types of math models
Step 3: From Optimizing to Satisficing

- **Course: How to find solutions to various types of math models**

- Herb Simon coined the term **satisficing** as a combination of the words satisfactory and optimizing

- **Satisficing (approximate solutions)**
  - Use Heuristic procedures to find suboptimal solutions
  - tendency of managers to seek a solution that is “good enough” for the problem at hand. Goals may be set to establish minimum satisfactory levels of performance in various areas, based perhaps on past levels of performance or on what the competition is achieving. If a solution is found that enables all these goals to be met, it is likely to be adopted without further ado. Such is the nature of satisficing.
Postoptimality Analysis: Sensitive Parameters

Step 3:

- **Analysis done after finding an optimal solution**
  - This analysis also is sometimes referred to as **what-if analysis** because it involves addressing some questions about what would happen to the optimal solution if different assumptions are made about future conditions.
  - These questions often are raised by the managers who will be making the ultimate decisions rather than by the OR team.

- **Postoptimality analysis involves conducting sensitivity analysis to determine which **parameters** of the model are most critical (the “sensitive parameters”) in determining the solution.**

- **The value assigned to a parameter commonly is just an estimate of some quantity (e.g., unit profit) whose exact value will become known only after the solution has been implemented.**
  - Therefore, after the sensitive parameters are identified, special attention is given to estimating each one more closely, or at least its range of likely values. One then seeks a solution that remains a particularly good one for all the various combinations of likely values of the sensitive parameters.
Step 4: Debugging the Model

• Developing a large mathematical model is analogous in some ways to developing a large computer program. When the first version of the computer program is completed, it inevitably contains many bugs.

• Post-optimality analysis is also key here.

• Reexamining the definition of the problem and comparing it with the model may help to reveal mistakes. (imagine an exam question)
  – It is also useful to make sure that all the mathematical expressions are dimensionally consistent in the units used. Additional insight into the validity of the model can sometimes be obtained by varying the values of the parameters and/or the decision variables and checking to see whether the output from the model behaves in a plausible manner.

• Retrospective test over held out data (just like ML!)
Step 5: Build Supporting Infrastructure

• Produce documentation
• Model and analysis reports
• Solution procedure

• Build supporting infrastructure for using such a model in production (accessing data, logging data, make decisions, etc.)
Step 6 Deploy (and AB Test)

- Run system in a production setting (e.g., in a data center)

- AB Test new system with incumbent system

- Refine model, repeat step 2-6 if necessary (live performance is not satisfactory)
Systems Modeling is inherently interactive and iterative.

1. Understand the domain and Define problems
2. Formulate solution/model
3. Derive solutions
4. Debug/Evaluate models
5. Build supporting infrastructure for models
6. Deploy System in the wild (and AB test)

6 Steps to OR Modeling in Practice
Case Studies in Digital Advertising
Advertising

- Advertising is a paid, one-way communication
  1. Deliver marketing messages and attract new customers
  2. To inform potential customers about products and services and how to obtain and use them.
  3. Branding → Direct action
     - Many advertisements are also designed to generate increased consumption of those products and services through the creation and reinforcement of brand image and brand loyalty (ads contain both factual information and persuasive messages).
  4. Use every major medium
     - To deliver these messages, including: television, radio, movies, magazines, newspapers, video games, the Internet, and billboards
Digital Advertising

• Online advertising is a form of advertising utilizing the Internet and World Wide Web in order to deliver marketing messages and attract customers [wikipedia.com]

• Advertising annoys people! Advertising works!
  – "Half the money I spend on advertising is wasted; the trouble is, I don't know which half." - John Wanamaker, father of modern advertising. [Credit assignment]
  – "I do not regard advertising as entertainment or an art form, but as a medium of information...“, “Ogilvy on Advertising” by David Ogilvy

• Goals of Online advertising

  A
  – Deliver/push an advertiser’s message with quantifiable measures of consumer interest

  A+P
  – Generate ROI for the advertiser and revenue for the publisher

  P+C
  – Enable ads as a medium of information (true in the case of search)!

Advertising makes up ~2% of US GDP

"Half the money I spend on advertising is wasted; the trouble is, I don't know which half." - John Wanamaker, father of modern advertising.

Less than 1% of all impressions lead to measurable ROI

Despite its problems (Attribution etc.)

- US GDP = $14.1 Trillion (Global $56 Trillion, $56 \times 10^{12})
- US Advertising Spend
  - ~$275 Billion across all media
    - (2% of GDP since the early 1900s)

- In 2008, Worldwide online advertising was $65B
  - I.e., about 10% of all ad spending across all media [IDC, 2008]
  - $23 Billion in US; $2 Billion in China; $2 Billion in Latin America;
  - $20B (Europe); and Russia accounted for $720 million
Purchase Funnel

Deliver marketing messages and attract customers and sell products/services

- **Awareness**
  - Targeting Demo, Geo and Content related websites

- **Consideration**
  - SEM: User who is searching for your product
  - Site Retargeting: Users who previously visited your website
  - users who is on your website

- **Conversion**
What marketers want via advertising?

- Deliver marketing messages and attract customers and sell products/services

<table>
<thead>
<tr>
<th>Goal</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduce: Reach</td>
<td>Media Planning</td>
</tr>
<tr>
<td>Influence: Brand</td>
<td>Ad Effectiveness (CTR, site visits)</td>
</tr>
<tr>
<td>Close</td>
<td>Marketing Effectiveness</td>
</tr>
<tr>
<td>Network Effect</td>
<td>Referrals/Advocacy</td>
</tr>
</tbody>
</table>

SEM: User who is searching for your product
User who previously searched for your product
Users who previously showed intent on interest related sites
Targeting Demo, Geo and Content related websites
Media Channels for Advertising

- Advertising online comes in all shapes and sizes and we run into it all the time be it through
  - Websearch
  - Reading the newspaper online
  - Paying the bills
  - Listening to music
  - Watching a video
  - Purchasing a book
  - Mobile device-based apps (phones, Tablet computers)
### Sponsored Search

<table>
<thead>
<tr>
<th>Search Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertising</td>
<td>Wikipedia, the free encyclopedia (4 visits - 3:08am) For content guidelines on advertising in Wikipedia articles, see en.wikipedia.org/wiki/Advertising.</td>
</tr>
<tr>
<td>Website Ad</td>
<td>Display Your Ad for Free &amp; Pay Only When Customers Respond to Your Ad! adwords.google.com</td>
</tr>
<tr>
<td>Marketing &amp; Advertising</td>
<td>Web, Print, Lead Generation, SEO Technology &amp; Startup Specialists <a href="http://www.glasscomply.com">www.glasscomply.com</a> San Francisco-Oakland-San Jose, CA</td>
</tr>
<tr>
<td>Free Online Advertising</td>
<td>Get Listed on Major Search Engines with a 30 Day Trial. No Risk! <a href="http://www.Yodle.com">www.Yodle.com</a></td>
</tr>
<tr>
<td>Facebook Advertising</td>
<td>Reach the exact audience you want with relevant targeted ads. <a href="http://www.facebook.com/ads/">www.facebook.com/ads/</a></td>
</tr>
<tr>
<td>Advertise in Your Area</td>
<td>Attract Local Customers to Your Business With Direct Advertising! <a href="http://www.valpax.com/advertise">www.valpax.com/advertise</a> San Francisco-Oakland-San Jose, CA</td>
</tr>
<tr>
<td>MySpace Advertising</td>
<td>Target ads by over 1,100 hobbies &amp; interests. Budgets as low as $5/day. Advertise.aMySpace.com</td>
</tr>
<tr>
<td>Buttons, Stickers &amp; More</td>
<td>Custom Creations w/your Message Service &amp; Quality you can Count on! <a href="http://www.EBlueCreations.com">www.EBlueCreations.com</a></td>
</tr>
</tbody>
</table>

---

**ISM 250: Stochastic Optimization in Info Sys and Tech** © 2011 James G. Shanahan James.Shanahan_AT_gmail.com
Display Ads

- Also called banner ads
- Served by web sites
- Similar to ‘Display’ or ‘Hoardings’ on road side
- Ads are targeted based on the demographics of the visitors of the web site
Contextual Ads:

Target based on text of page

- Served by the web site to its visitors
- Ad network selects ads that are highly related to the content of the web page
Contextual Advertising on Webpages

For standards see IAB
http://www.iab.net/standards/adunits.asp
House Ads at AMEX
House Ads at Amazon
OA is cavalier! : business models; ad placement; e:b wants to be online
Advertising: a supply-demand marketplace

**DEMAND**
Advertiser wishes to reach consumers

**SUPPLY**
Ad Slots for sale

Formal Relationship

Consumers
Supply

DEMAND
Advertiser wishes to reach consumers

Ads

Agency

ADVERTISER

Agency

SUPPLY

• Sponsored Search
• Contextual
• Display
• Classified
• Email
• Social
• Mobile

Publish

Consumers
From Mad Men To Wall Street and beyond!

- Set in New York City, *Mad Men* begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.

<table>
<thead>
<tr>
<th>Banner</th>
<th>Click+Data</th>
<th>Personal</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Intensive</td>
<td>Technology Data Driven</td>
<td>Data</td>
<td>Social</td>
</tr>
<tr>
<td>Lots of guess work</td>
<td>Forward Market Spot Markets</td>
<td>Personalization</td>
<td></td>
</tr>
<tr>
<td>Forward Market</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2007

1st Generation

2nd Generation

3rd Generation

Advertisers still in broadcast mode

YoY: Double digit growth
Update on DSP Bidding Competition

沈学华
x@ipinyou.com
3 Milestones + 1,000,000 Grand Prize
Happy New Year 2014

© Can Stock Photo - cs18528917
Maximize \#clicks + N \times \#conversions
Subject to the fixed budget
Dropbox:
https://www.dropbox.com/sh/xolf5thu8jsbmfu/kBrAsSxtAN

百度网盘:
http://pan.baidu.com/share/link?shareid=374646&uk=3037373637
7.5G

bid: 13.6M
imp: 9.2M
clk: 7.5K
conv: 72
Any question sent to
dsp-competition@ipinyou.com
2012 KDD Cup

Predict the click-through rate of ads given the query and user information.

TASK 2 DESCRIPTION

Search advertising has been one of the major revenue sources of the Internet industry for years. A key technology behind search advertising is to predict the click-through rate (pCTR) of ads, as the economic model behind search advertising requires pCTR values to rank ads and to price clicks. In this task, given the training instances derived from session logs of the Tencent proprietary search engine, soso.com, participants are expected to accurately predict the pCTR of ads in the testing instances.

TRAINING DATA FILE

The training data file is a text file, where each line is a training instance derived from search session log messages. To understand the training data, let us begin with a description of search sessions.

A search session refers to an interaction between a user and the search engine. It contains the following ingredients: the user, the query issued by the user, some ads returned by the
Continuing the Learning Process with Your OR Courseware
On Windows 8 (via Parallels on Mac): Press Command Key for a couple of Seconds to bring up Windows VM
Then Right Click to bring up APPs
Make sure Java is installed

Free Java Download

Download Java for your desktop computer now!

Version 7 Update 10

Free Java Download

Why download Java?

Java technology allows you to work and play in a secure computing environment.

(1) IORTutorial (interactive) and (2) OR Tutor
IOR Tutor

Press the Menu Help button to go through a general introduction to this software.

If you wish to begin work now instead, choose an Area from the Area menu.

* Copyright 2004 by McGraw-Hill, Inc. All rights reserved.
Interactive Operations Research Tutorial

- Enter or Revise a General Linear Programming Model
- Solve Automatically by the Interior Point Algorithm
- Set Up for the Simplex Method--Interactive only
- Solve Interactively by the Simplex Method
- Sensitivity Analysis
- Modified Simplex Method
- Solve Automatically by the Simplex Method
- Graphical Method and Sensitivity Analysis
- Solve Interactively by the Graphical Method

Procedure from the Procedure menu to start. For some of the procedures, an OR Tutor to help prepare you to go through the procedures in order. The first procedure must precede the second, second must precede the third, and the fourth must precede the fifth. The sixth procedure is the Programming Area, and instructions for this procedure are given under the Area menu. Within the Programming Area, select the Area menu to work.
Objective Function: $\text{Max} \quad Z = 5x_1 + 3x_2$ 

Constraint 

Note:
The number of constraints can't be more than 5
Subject to:
$x_1 \geq 0, \ x_2 \geq 0$

(1) $5x_1 \leq 9$
(2) $3x_1 + 4x_2 \leq 12$

$x_2$

Optimum

$Z = 13.95$ with $x_1 = 1.8, \ x_2 = 1.65$

After the optimal solution is reached, you can do sensitivity analysis by pressing the button below.

Sensitivity Analysis

A: General Linear Programming  P: Graphical Method and Sensitivity Analysis  O: Tabular Form
Demo for LP and Sensitivity Analysis

- Graphical Method and Sensitivity Analysis
- Interactive Operations Research Tutorial
The Graphical Method - Formulation

Using money units of **thousands of dollars** and time units of **hours**, the data for Leo's problem can be summarized as follows.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Resource Usage per Share of Each Venture</th>
<th>Amount of Resource Available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Venture 1</td>
<td>Venture 2</td>
</tr>
<tr>
<td>Money</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>Time</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Savings per share</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

Since Leo must decide how many shares of each venture to take, the **decision variables** are

\[ x_1 = \text{number of shares of Venture 1}, \]
\[ x_2 = \text{number of shares of Venture 2}. \]

He wants to maximize his savings, so the value of the objective function is

\[ Z = \text{number of money units (thousands of dollars) added to savings}. \]

The linear programming model for Leo's problem now can be formulated from the data in the table as follows.

Maximize \[ Z = 20x_1 + 10x_2 \]

subject to

\[ x_1 - x_2 \leq 1 \]
\[ 3x_1 + x_2 \leq 7 \]

and \[ x_1 \geq 0, \; x_2 \geq 0. \]

Boot up Explorer in Parallels and then type in: C:\Program Files (x86)\OR Tutor\index.html
Another graphical example in OR Tutor

The Graphical Method - Optimal Solution

Maximize \( Z = 20x_1 + 10x_2 \)

subject to \( x_1 - x_2 \leq 1 \) \hspace{1em} \text{(money constraint)}

\( 3x_1 + x_2 \leq 7 \) \hspace{1em} \text{(time constraint)}

\text{and} \hspace{1em} x_1 \geq 0, x_2 \geq 0.

Optimal Solution:
Optimal point on graph: (0, 7).

Resulting optimal solution: \( x_1 = 0, x_2 = 7 \).

Optimal value of \( Z \): \( Z = 70 \).

Remember that \( x_1 \) and \( x_2 \) represent Leo's number of shares of Ventures 1 and 2 respectively. Also, \( Z \) is measured in units of thousands of dollars. Therefore, our conclusion is that Leo should take 7 shares of Venture 2, but none of Venture 1. This strategy will enable Leo to add $70,000 dollars to his savings before returning to school to earn an MBA degree.

This concludes the demonstration. See the OR Tutor menu (to the left) for other demonstrations or close the browser window to exit OR Tutor.
OR Tutor: supplement examples in the book

The Graphical Method - A supplement example in the book

Using money units of thousands, the data for Leo's problem can be summarized as follows.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Amount of Resource Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money</td>
<td>1</td>
</tr>
<tr>
<td>Time</td>
<td>7</td>
</tr>
<tr>
<td>Savings</td>
<td>5</td>
</tr>
</tbody>
</table>

Since Leo must decide how much of each product to produce, the decision variables are

\[ x_1, x_2 = \text{amount of Product } 1, \text{ Product } 2 \]

He wants to maximize his savings, so the objective function is

\[ Z = 3x_1 + 5x_2 \]

The linear programming model is

\[ \text{Maximize } Z = 3x_1 + 5x_2 \]

subject to

\[ 2x_1 + 3x_2 \leq 1 \]

\[ x_1 \geq 0, x_2 \geq 0 \]
Other systems

- Solve a variety of OR models using
  - The Excel Solver
  - LINGO/LINDO
  - MPL/CPLEX
  - R (optim)
      - `install.packages("boot")`
      - `library("boot")`
install.packages("boot")
library("boot")
# This example is taken from Exercise 7.5 of Gill, Murray,
# and Wright (1991).
enj <- c(200, 6000, 3000, -200)
fat <- c(800, 6000, 1000, 400)
vitx <- c(50, 3, 150, 100)
vity <- c(10, 10, 75, 100)
vitz <- c(150, 35, 75, 5)
simplex(a = enj, A1 = fat, b1 = 13800, A2 = rbind(vitx, vity, vitz),
       b2 = c(600, 300, 550), maxi = TRUE)
package 'boot' was built under R version 2.14.1

> # This example is taken from Exercise 7.5 of Gill, Murray, 
> # and Wright (1991).
> enj <- c(200, 6000, 3000, -200)
> fat <- c(800, 6000, 1000, 400)
> vitx <- c(50, 3, 150, 100)
> vity <- c(10, 10, 75, 100)
> vitz <- c(150, 35, 75, 5)
> simplex(a = enj, A1 = fat, b1 = 13800, A2 = rbind(vitx, vity, vitz), 
> +   b2 = c(600, 300, 550), maxi = TRUE)

Linear Programming Results

Call : simplex(a = enj, A1 = fat, b1 = 13800, A2 = rbind(vitx, vity, 
   vitz), b2 = c(600, 300, 550), maxi = TRUE)

Maximization Problem with Objective Function Coefficients

   x1  x2  x3  x4  
200 6000 3000 -200

Optimal solution has the following values

   x1  x2  x3  x4  
0.0 0.0 13.8 0.0

The optimal value of the objective function is 41400.
• End of Lecture 1