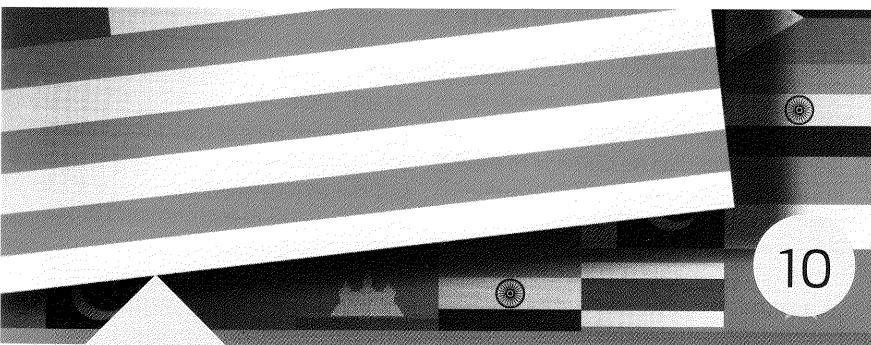


Cover story



10

Nation in a dilemma

Although they contribute significantly to Malaysia's economy, migrant workers continue to be socially and legally ignored.

Editorial / pg 4

Letter to the editor / pg 5

Mission statement / pg 7

SERI revamps its magazine in response to Penang's new political, economic and cultural scenario.

George Town Festival / pg 40

The George Town World Heritage Office is set to win the hearts and minds of tourists and Penangites.



Laying the groundwork for innovative growth / pg 42

Malaysia can still raise its productivity and compete in the global economy.

Events / pg 63

JUST A MINUTE

Voter registration / pg 8

The race to register new voters is on!

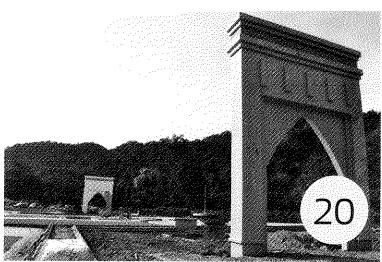
PHOTO ESSAY

Lights, camera... Penang State Assembly! / pg 24

We go behind the scenes at the 12th Penang Legislative Assembly.



Feature

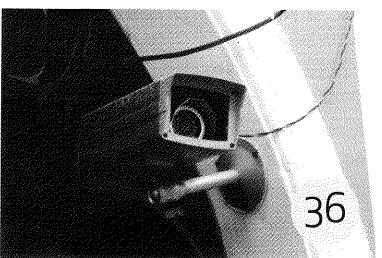


20

Gardens going to seed

The Penang Botanic Gardens have been making the news lately – all for the wrong reasons.

States of reform



36

Safety in the cities

Crime rates are exploding throughout the nation, and Malaysians continue to live in fear.

Utter economics



54

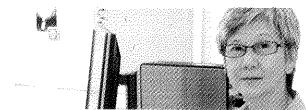
Unravelling the figures

How to read the 2009 Bank Negara Report.

PROFILE

An eminent Malaysian academic abroad / pg 30

This month, we profile an internationally-respected anthropologist, and chat with her about Malay society in the 1970s.



VOICE IN THE CROWD

Butterworth remains the Ugly Duckling / pg 34

There is hope yet for the long-ignored Butterworth.



PEOPLE ECONOMICS

The New Economic Model - Strategic Policy Directions / pg 46

A quick summary of the New Economic Model.

PENANG PALETTE

Birdmen of Penang / pg 48

Penang's bird artists lament the decrease in the island's bird population.



STATISTICS / pg 60

BIG DATA

10 Guest Editors' Introduction

Francis J. Alexander, Adolfy Hoisie, and Alexander Szalay

14 Data-Intensive Science in the Department of Energy: Case Studies and Future Challenges

James P. Ahrens, Bruce Hendrickson, Gabrielle Long, Steve Miller, Robert Ross, and Dean Williams

Given its leading role in high-performance computing for modeling and simulation and its many experimental facilities, the US Department of Energy has a tremendous need for data-intensive science. Locating the challenges and commonalities among three case studies illuminates, in detail, the technical challenges involved in realizing data-intensive science.

25 Data-Intensive Scalable Computing for Scientific Applications

Randal E. Bryant

Increasingly, scientific computing applications must accumulate and manage massive datasets, as well as perform sophisticated computations over these data. Such applications call for data-intensive scalable computer (DISC) systems, which differ in fundamental ways from existing high-performance computing systems.

34 Extreme Data-Intensive Scientific Computing

Alexander S. Szalay

Scientific computing increasingly involves massive data; in astronomy, observations and numerical simulations are on the verge of generating petabytes. This new, data-centric computing requires a new look at computing architectures and strategies. Using Amdahl's law to characterize architectures and workloads, it's possible to use existing commodity parts to build systems that approach an ideal Amdahl machine.

ALSO IN THIS ISSUE

42 Massively Parallel Neural Signal Processing on a Many-Core Platform

Dan Chen, Lizhe Wang, Gaoxiang Ouyang, and Xiaoli Li

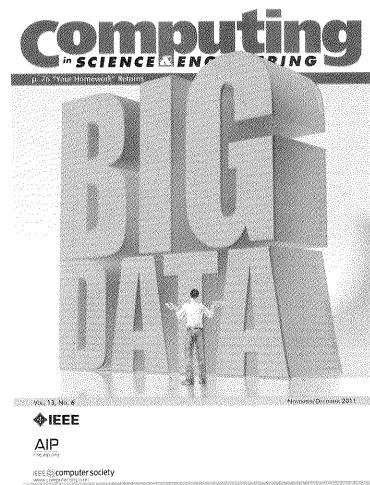
Although the ensemble empirical mode decomposition (EEMD) method and Hilbert-Huang transform (HHT) offer an unrivaled opportunity to understand neural signals, the EEMD algorithm's complexity and neural signals' massive size have hampered EEMD application. However, a new approach using a many-core platform has proven both efficient and effective for massively parallel neural signal processing.

52 Visualizing 3D Earthquake Simulation Data

Cheng-Kai Chen, Chris Ho, Carlos Correa, Kwan-Liu Ma, and Ahmed Elgamal

As exemplified in a state-of-the-art bridge-foundation-ground model simulation, a suite of new visualization techniques lets scientists study seismic waves and interactively investigate and explore their data. In so doing, the techniques further scientific understanding and thus facilitate the development of new methods to protect real-world infrastructures against otherwise devastating earthquakes.

For more information on these and other computing topics, please visit the IEEE Computer Society Digital Library at www.computer.org/publications/dlib/.



Cover illustration: Giacomo Marchesi
www.giacomomarchesi.com

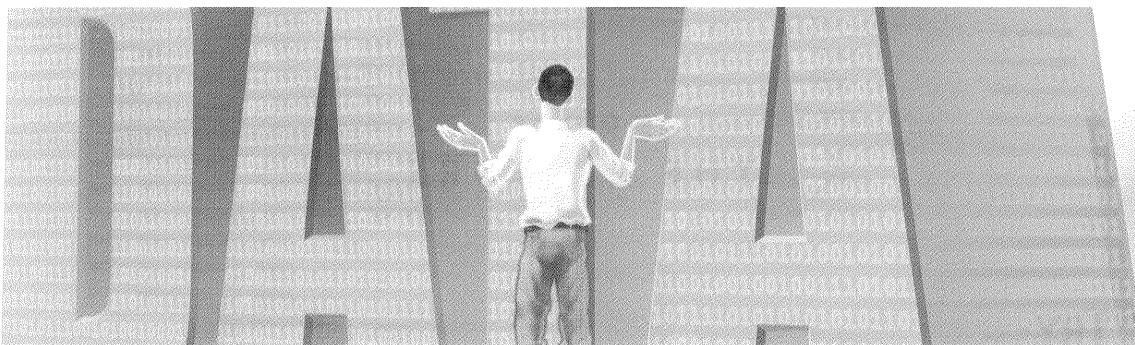
STATEMENT OF PURPOSE

Computing in Science & Engineering aims to support and promote the emerging discipline of computational science and engineering and to foster the use of computers and computational techniques in scientific research and education. Every issue contains broad-interest theme articles, departments, news reports, and editorial comment. Collateral materials such as source code are made available electronically over the Internet. The intended audience comprises physical scientists, engineers, mathematicians, and others who would benefit from computational methodologies. All articles and technical notes in *CISE* are peer-reviewed.



Printed with
inks containing
soy and/or
vegetable oils





COLUMNS

4 From the Editors

Just Try
Isabel Beichl

96 The Last Word

It's More Fun to Compose
Charles Day

DEPARTMENTS

6 Books

No Room for Filler: Large Content, Small Book
David Szurley

Teaching Yourself CUDA: Learn to Walk While
You Run
Manuel Tiglio

64 Visualization Corner

General-Relativistic Visualization
Thomas Müller and Daniel Weiskopf

72 Education

The Numerical Tours of Signal Processing, Part 3:
Image and Surface Restoration
Gabriel Peyré

76 Your Homework

NASA Computational Case Study
Characterizing Moving Particles
Nargess Memarsadeghi and Brent J. Bos

79 Novel Architectures

SCC: A Flexible Architecture for Many-Core
Platform Research
*Matthias Gries, Ulrich Hoffmann, Michael Konow,
and Michael Riepen*

84 Computer Simulations

Coupled Time Series Analysis: Methods
and Applications
*Gholamreza Jafari, Amir Hossein Shirazi,
Ali Namaki, and Reza Raei*

90 Scientific Programming

Using the Data Access Protocol with IDL
Michael D. Gallo

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *Computing in Science & Engineering* does not necessarily constitute endorsement by the IEEE, the IEEE Computer Society, or the AIP. All submissions are subject to editing for style, clarity, and length. IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

Circulation: *Computing in Science & Engineering* (ISSN 1521-9615) is published bimonthly by the AIP and the IEEE Computer Society. IEEE Headquarters, Three Park Ave., 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Cir., PO Box 3014, Los Alamitos, CA 90720-1314, phone +1 714 821 8380; IEEE Computer Society Headquarters, 2001 L St., Ste. 700, Washington, D.C. 20036; AIP Circulation and Fulfillment Department, 1NO1, 2 Huntington Quadrangle, Melville, NY, 11747-4502. 2011 annual subscription rates: US\$49 for Computer Society members (print plus online), \$78 (sister society), and \$625 (individual nonmember). For AIP society members, 2011 annual subscription rates are \$49 (print plus online). For more information on other subscription prices, see www.computer.org/subscribe/ or https://www.aip.org/forms/journal_catalog/order_form_fs.html. Computer Society back issues cost \$20 for members, \$111 for nonmembers; AIP back issues cost \$22 for members. **Reuse Rights and Reprint Permissions:** Educational or personal use of this material is permitted without fee, provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of IEEE-copyrighted material on their own web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first screen of the posted copy. An accepted manuscript is a version that has been revised by the author to incorporate review suggestions, but not the published version with copy-editing, proofreading and formatting added by IEEE. For more information, please go to: http://www.ieee.org/publications_standards/publications/rights/paperversionpolicy.html. Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee.org. Copyright © 2011 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Dr., Danvers, MA 01923. Postmaster: Send undelivered copies and address changes to *Computing in Science & Engineering*, 445 Hoes Ln., Piscataway, NJ 08855. Periodicals postage paid at New York, NY, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in the USA.

RESOURCES

8 AIP Membership Information

9 IEEE Computer Society Information

75 Advertiser Index

CiSE's Annual Index for 2011 is available online at www.computer.org/cise/11index

IN THIS ISSUE

12 XML Management for Bioinformatics Applications

Lena Strömbäck and Juliana Freire

Scientific exploration has become a data-intensive process, and increasing amounts of data need to be stored, analyzed, and shared. XML can help address these needs. As concrete examples from systems biology show, native XML storage can be combined with traditional relational databases to offer an effective, usable solution for storing scientific data.

24 Survey: Real-Time Tumor Motion Prediction for Image-Guided Radiation Treatment

Poonam S. Verma, Huanmei Wu, Mark P. Langer, Indra J. Das, and George Sandison
Tumor motion caused by patient breathing creates challenges for accurate radiation dose delivery to a tumor while sparing healthy tissues. Image-guided radiation therapy (IGRT) helps, but there's a lag time between tumor position acquisition and dose delivered to that position. An efficient and accurate predictive model is thus an essential requirement for IGRT success.

36 Climate Change Modeling: Computational Opportunities and Challenges

Dali Wang, Wilfred M. Post, and Bruce E. Wilson

High-fidelity climate models are the workhorses of modern climate change sciences. In this article, the authors focus on several computational issues associated with climate change modeling, covering simulation methodologies, temporal and spatial modeling restrictions, the role of high-end computing, as well as the importance of data-driven regional climate impact modeling.

43 ScalaLab: An Effective Scala-Based Scientific Programming Environment for Java

Stergios Papadimitriou, Konstantinos Terzidis, Seferina Mavroudi, and Spiridon Likothanassis

Scala offers many benefits for constructing scientific programming environments. Extending Scala with Matlab-like constructs enabled the creation of ScalaSci, a compiled mathematical scripting framework, and ScalaLab, an efficient integrated scientific programming environment. ScalaLab offers an interesting open source alternative to commercial packages, especially for the scientific community familiar with Java.

56 Coupling Advanced Modeling and Visualization to Improve High-Impact Tropical Weather Prediction

Bo-Wen Shen, Wei-Kuo Tao, and Bryan Green

To meet the goals of extreme weather event warning, this approach couples a modeling and visualization system that integrates existing NASA technologies and improves the modeling system's parallel scalability to take advantage of petascale supercomputers. It also streamlines the data flow for fast processing and 3D visualizations, and develops visualization modules to fuse NASA satellite data.

For more information on these and other computing topics, please visit the IEEE Computer Society Digital Library at www.computer.org/publications/dlib/.

computing

in SCIENCE & ENGINEERING



Cover illustration: Giacomo Marchesi
www.giacomomarchesi.com

STATEMENT OF PURPOSE

Computing in Science & Engineering aims to support and promote the emerging discipline of computational science and engineering and to foster the use of computers and computational techniques in scientific research and education. Every issue contains broad-interest theme articles, departments, news reports, and editorial comment. Collateral materials such as source code are made available electronically over the Internet. The intended audience comprises physical scientists, engineers, mathematicians, and others who would benefit from computational methodologies. All articles and technical notes in *CISE* are peer-reviewed.



Printed with
inks containing
soy and/or
vegetable oils





COLUMNS

4 From the Editors

Top of the Charts
Isabel Beichl

96 The Last Word

Standards Rule OK
Charles Day

DEPARTMENTS

5 Books

The GPU Revolution at Work
Martin Weigel

7 News

Software Engineering for Scientists
Diane Kelly, Spencer Smith, and Nicholas Meng

68 Education

The Numerical Tours of Signal Processing Part 2:
Multiscale Processings
Gabriel Peyré

72 Visualization Corner

Streaming-Enabled Parallel Data Flow
Framework in the Visualization ToolKit
*Huy T. Vo, João L.D. Comba, Berk Geveci,
and Cláudio T. Silva*

84 Scientific Programming

Moving Academic Department Functions to
Social Networks and Clouds: Initial Experiences
*George K. Thiruvathukal, Konstantin Läufer,
and David Dennis*

90 Novel Architectures

Keeneland: Bringing Heterogeneous
GPU Computing to the Computational
Science Community
*Jeffrey S. Vetter, Richard Glassbrook,
Jack Dongarra, Karsten Schwan, Bruce Loftis,
Stephen McNally, Jeremy Meredith,
James Rogers, Philip Roth, Kyle Spafford,
and Sudhakar Yalamanchili*

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *Computing in Science & Engineering* does not necessarily constitute endorsement by the IEEE, the IEEE Computer Society, or the AIP. All submissions are subject to editing for style, clarity, and length. IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

Circulation: *Computing in Science & Engineering* (ISSN 1521-9615) is published bimonthly by the AIP and the IEEE Computer Society. IEEE Headquarters, Three Park Ave., 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Cir., PO Box 3014, Los Alamitos, CA 90720-1314, phone +1 714 821 8380; IEEE Computer Society Headquarters, 2001 L St., Ste. 700, Washington, D.C. 20036; AIP Circulation and Fulfillment Department, 1NO1, 2 Huntington Quadrangle, Melville, NY, 11747-4502. 2011 annual subscription rates: US\$49 for Computer Society members (print plus online), \$78 (sister society), and \$625 (individual nonmember). For AIP society members, 2011 annual subscription rates are \$49 (print plus online). For more information on other subscription prices, see www.computer.org/subscribe/ or https://www.aip.org/forms/journal_catalog/order_form_fs.html. Computer Society back issues cost \$20 for members, \$111 for nonmembers; AIP back issues cost \$22 for members. **Reuse Rights and Reprint Permissions:** Educational or personal use of this material is permitted without fee, provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of IEEE-copyrighted material on their own web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first screen of the posted copy. An accepted manuscript is a version that has been revised by the author to incorporate review suggestions, but not the published version with copy-editing, proofreading and formatting added by IEEE. For more information, please go to: http://www.ieee.org/publications_standards/publications/rights/paperversionpolicy.html. Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee.org. Copyright © 2011 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Dr., Danvers, MA 01923. Postmaster: Send undelivered copies and address changes to *Computing in Science & Engineering*, 445 Hoes Ln., Piscataway, NJ 08855. Periodicals postage paid at New York, NY, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in the USA.

RESOURCES

11 IEEE Computer Society Information

22 AIP Membership Information

55 Advertiser Index

The Capital Asset Pricing Model

16.1 Introduction to the CAPM

The *CAPM (capital asset pricing model)* has a variety of uses. It provides a theoretical justification for the widespread practice of passive investing by holding *index funds*.¹ The CAPM can provide estimates of expected rates of return on individual investments and can establish “fair” rates of return on invested capital in regulated firms or in firms working on a cost-plus basis.²

The CAPM starts with the question, what would be the risk premiums on securities if the following assumptions were true?

1. The market prices are “in equilibrium.” In particular, for each asset, supply equals demand.
2. Everyone has the same forecasts of expected returns and risks.
3. All investors choose portfolios optimally according to the principles of efficient diversification discussed in Chapter 11. This implies that everyone holds a tangency portfolio of risky assets as well as the risk-free asset.
4. The market rewards people for assuming unavoidable risk, but there is no reward for needless risks due to inefficient portfolio selection. Therefore, the risk premium on a single security is not due to its “standalone” risk, but rather to its contribution to the risk of the tangency portfolio. The various components of risk are discussed in Section 16.4.

Assumption 3 implies that the market portfolio is equal to the tangency portfolio. Therefore, a broad index fund that mimics the market portfolio can be used as an approximation to the tangency portfolio.

The validity of the CAPM can only be guaranteed if all of these assumptions are true, and certainly no one believes that any of them are exactly true.

¹ An index fund holds the same portfolio as some index. For example, an S&P 500 index fund holds all 500 stocks on the S&P 500 in the same proportions as in the index. Some funds do not replicate an index exactly, but are designed to track the index, for instance, by being cointegrated with the index.

² See Bodie and Merton (2000).

Risk Management

19.1 The Need for Risk Management

The financial world has always been risky, and financial innovations such as the development of derivatives markets and the packaging of mortgages have now made risk management more important than ever but also more difficult.

There are many different types of risk. *Market risk* is due to changes in prices. *Credit risk* is the danger that a counterparty does not meet contractual obligations, for example, that interest or principal on a bond is not paid. *Liquidity risk* is the potential extra cost of liquidating a position because buyers are difficult to locate. *Operational risk* is due to fraud, mismanagement, human errors, and similar problems.

Early attempts to measure risk such as duration analysis, discussed in Section 3.8.1 and used to estimate the market risk of fixed income securities, were somewhat primitive and of only limited applicability. In contrast, value-at-risk (VaR) and expected shortfall (ES) are widely used because they can be applied to all types of risks and securities, including complex portfolios.

VaR uses two parameters, the time horizon and the confidence level, which are denoted by T and $1 - \alpha$, respectively. Given these, the VaR is a bound such that the loss over the horizon is less than this bound with probability equal to the confidence coefficient. For example, if the horizon is one week, the confidence coefficient is 99% (so $\alpha = 0.01$), and the VaR is \$5 million, then there is only a 1% chance of a loss exceeding \$5 million over the next week. We sometimes use the notation $\text{VaR}(\alpha)$ or $\text{Var}(\alpha, T)$ to indicate the dependence of VaR on α or on both α and the horizon T . Usually, $\text{VaR}(\alpha)$ is used with T being understood.

If \mathcal{L} is the loss over the holding period T , then $\text{VaR}(\alpha)$ is the α th upper quantile of \mathcal{L} . Equivalently, if $\mathcal{R} = -\mathcal{L}$ is the revenue, then $\text{VaR}(\alpha)$ is minus the α th quantile of \mathcal{R} . For continuous loss distributions, $\text{VaR}(\alpha)$ solves

$$P\{\mathcal{L} > \text{VaR}(\alpha)\} = P\{\mathcal{L} \geq \text{VaR}(\alpha)\} = \alpha, \quad (19.1)$$

and for any loss distribution, continuous or not,



Decision-making with uncertain data: Bayesian linear programming approach

A. P. TCHANGANI

EPA-LGP, Dpt. GEII, IUT de Tarbes, Université Toulouse III 1, rue Lautréamont, 65016 Tarbes Cedex, France

EPA-LGP, 47, Av. d'Azereix, B.P. 1629, 65016 Tarbes, France

E-mail: tchangani@geii.iut-tarbes.fr

Received May 2002 and accepted March 2003

This paper deals with decision making in a real time optimization context under uncertain data by linking Bayesian networks (BN) techniques (for uncertainties modeling) and linear programming (LP, for optimization scheme) into a single framework. It is supposed that some external events sensed in real time are susceptible to give relevant information about data. BN consists in graphical representation of probabilistic relationship between variables of a knowledge system and so permit to take into account uncertainty in an expert system by bringing together the classical artificial intelligence (AI) approach and statistics approach. They will be used to estimate numerical values of parameters subjected to the influence of random events for a linear programming program that perform optimization process in order to select optimal values of decision variables of a certain real time decision-making system.

Keywords: Linear programming, Bayesian networks, decision making, decision support system, real time optimization

1. Introduction and statement of the problem

1.1. Decision-making by mathematical programming

Any organization, from a single person to a multi-national company through small companies, factories, etc., make always decisions (investment, resources acquisition, production planning, etc.). Decisions are made by choosing numerical values of some free variables called decision variables or control variables (production level for a given period, stock level, investment budget level, etc.) in order to optimize (minimize or maximize) some objectives (maximize profit, minimize cost, minimize risk, etc.) when respecting some constraints due to necessarily limitation of available resources (budget limitation, lack of qualified manpower, etc.) and some technical specifications. For instance, a production manager of a given company must establish a production plan by choosing

the level of each item to be produced for a given period (a day, a week or a month) using informations given by commercial and/or marketing department; a project manager must establish the planning of its crews respecting social and labor laws constraints; a logistic engineer must do the planning of reception dates of the resources to be used in a production process according to some requirements; and so on. These problems are well formulated when all the parameters are specified by the so-called mathematical programming (also known as optimization) problem in the literature. Basically, a mathematical programming problem consists of the following:

$$\left\{ \begin{array}{l} \min / \max_x f(x) \\ G(x) \leq 0 \\ x \in \mathcal{D} \subseteq \mathbb{R}^n \end{array} \right. \quad (1)$$

where $x \in \mathbb{R}^n$ is decision variables vector, $G(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is constraints vector, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$

Chapter 17

Portfolio Theory, CAPM and Performance Measures

Luis Ferruz, Fernando Gómez-Bezares, and María Vargas

Abstract This chapter is focused on the “Portfolio Theory” created by Markowitz. This theory has the objective of finding the optimum portfolio for investors; that is, that which gives tangency between an indifference curve and the efficient frontier. In this chapter, the mathematics of this model is developed.

The CAPM, based on this theory, gives the expected return on an asset depending on the systematic risk of the asset. This model detects underpriced and overpriced assets. The critics expressed against the model and their application possibilities are also analyzed.

Finally, the chapter centers on performance measures related to portfolio theory (classic indices, derivative indices and new approaches) and on the performance persistence phenomenon employing the aforementioned indices, including an empirical example.

Keywords CAPM • Performance measures • Penalized internal rate of return index (PIRR) • Penalized internal rate of return for beta index (PIRR for Beta) • Portfolio theory

17.1 Portfolio Theory and CAPM: Foundations and Current Application

17.1.1 Introduction

Portfolio theory has become highly developed and has strong theoretical support, making it essential in the toolbox of any financial expert.

In 1990, three American economists shared the Nobel Prize for Economics: Harry Markowitz, Merton Miller, and

William Sharpe. Markowitz is considered the creator of portfolio theory. He began this field of research with his work of 1952, making major additions with his works of 1959 and 1987. Sharpe attempted to simplify the model of his mentor, Markowitz, with his work of 1963, based on which he created the capital asset pricing model (CAPM) in 1964. Miller has also worked on portfolio topics.

Another Nobel Prize winner for Economics, James Tobin, contributed to the advance of portfolio theory with his work in 1958. Merton,¹ Black and Fama are others who have made great advances in the field.

A portfolio is a combination of assets, such as a group of stocks that are diversified to some extent. A portfolio is characterized by the return it generates over a given holding period. Portfolio theory describes the process by which investors seek the best possible portfolio in terms of the tradeoff of risk for return. In this context, risk refers to the situation where the investor is uncertain about which outcomes may eventuate, but can assess the probability with which each outcome may arise. In a risky environment, the choice of an optimal portfolio will depend on the distribution of returns, as well as the preferences of the investor – his or her appetite for risk.

It is clear that all decision makers have their own personal preferences, but it is possible to posit some generalizations. First, let us suppose that all decision makers prefer more wealth to less and, second, that decision makers are risk averse.

We define risk as the variability of the results of an activity; this variability can be measured by standard deviation or by variance. If we assume, as is customary, the marginal utility of an individual as a decreasing function of wealth, then the least risky activity is preferred, given the same average results.

Although the risk aversion behavior of individuals has been questioned, the literature is based on the assumption

L. Ferruz (✉) and M. Vargas
University of Zaragoza, Zaragoza, Spain
e-mail: lferruz@unizar.es

F. Gómez-Bezares
University of Deusto, Bilbao, Spain

¹ Merton also won a Nobel Prize in 1997 along with Scholes.

Factor Models and Principal Components

17.1 Dimension Reduction

High-dimensional data can be challenging to analyze. They are difficult to visualize, need extensive computer resources, and often require special statistical methodology. Fortunately, in many practical applications, high-dimensional data have most of their variation in a lower-dimensional space that can be found using *dimension reduction techniques*. There are many methods designed for dimension reduction, and in this chapter we will study two closely related techniques, *factor analysis* and *principal components analysis*, often called *PCA*.

PCA finds structure in the covariance or correlation matrix and uses this structure to locate low-dimensional subspaces containing most of the variation in the data.

Factor analysis explains returns with a smaller number of fundamental variables called *factors* or *risk factors*. Factor analysis models can be classified by the types of variables used as factors, macroeconomic or fundamental, and by the estimation technique, time series regression, cross-sectional regression, or statistical factor analysis.

17.2 Principal Components Analysis

PCA starts with a sample $\mathbf{Y}_i = (Y_{i,1}, \dots, Y_{i,d})$, $i = 1, \dots, n$, of d -dimensional random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. One goal of PCA is finding “structure” in $\boldsymbol{\Sigma}$.

We will start with a simple example that illustrates the main idea. Suppose that $\mathbf{Y}_i = \boldsymbol{\mu} + W_i \mathbf{o}$, where W_1, \dots, W_n are i.i.d. mean-zero random variables and \mathbf{o} is some fixed vector, which can be taken to have norm 1. The \mathbf{Y}_i lie on the line that passes through $\boldsymbol{\mu}$ and is in the direction given by \mathbf{o} , so that all variation among the mean-centered vectors $\mathbf{Y}_i - \boldsymbol{\mu}$ is in the one-dimensional space spanned by \mathbf{o} . Also, the covariance matrix of \mathbf{Y}_i is



Chapter 30

Itô's Calculus and the Derivation of the Black–Scholes Option-Pricing Model

George Chalamandaris and A.G. Malliaris

Abstract The purpose of this paper is to develop certain relatively recent mathematical discoveries known generally as *stochastic calculus*, or more specifically as *Itô's Calculus* and to also illustrate their application in the pricing of options. The mathematical methods of stochastic calculus are illustrated in alternative derivations of the celebrated Black–Scholes–Merton model. The topic is motivated by a desire to provide an *intuitive* understanding of certain probabilistic methods that have found significant use in financial economics.

Keywords Stochastic calculus • Itô's Lemma • Options pricing

30.1 Introduction

The purpose of this chapter is to develop certain relatively recent mathematical discoveries known generally as *stochastic calculus* (or more specifically as *Itô's Calculus*), describe some more modern methods, and illustrate their application in the pricing of options. The topic is motivated by a desire to provide an *intuitive* understanding of certain probabilistic methods that have found significant use in financial economics. A rigorous presentation of the same ideas is presented briefly in Malliaris (1983), and more extensively in Malliaris and Brock (1982), Duffie (1996) or Karatzas and Shreve (2005). Lamberton and Lapeyre (2007), Oksendal (2007), and Neftci (2000) are also useful books on stochastic calculus.

Itô's Calculus was prompted by purely mathematical questions originating in N. Wiener's work in 1923 on

G. Chalamandaris
Athens University of Economics and Business, Athens, Greece
e-mail: gchalamandar@auer.gr

A.G. Malliaris (✉)
School of Business Administration, Loyola University Chicago,
Chicago, IL, USA
e-mail: tmallia@luc.edu

stochastic integrals and was developed by the Japanese probabilist Kiyosi Itô during 1944–1951. Two decades later economists such as Merton (1973) and Black and Scholes (1973) started using Itô's stochastic differential equation to describe the behavior of asset prices. The results of Black–Scholes–Merton were soon generalized in the larger framework of the fundamental theorem of asset pricing following the work of Harrison and Kreps (1979) and Duffie and Huang (1985–1986). According to this theorem the absence of an arbitrage opportunity was linked to the notion of a Martingale measure and process. Because stochastic calculus is now used regularly by financial economists, some attention must be given to its mathematical meaning, its appropriateness and applications in economic modeling, and its applications to finance.

30.2 The ITÔ Process and Financial Modeling

Stochastic calculus is the mathematical treatment of random change in continuous time, unlike ordinary calculus, which deals with deterministic change.

This section defines intuitively the Brownian motion (or Wiener process), the Itô integral and equation, and discusses its appropriateness to financial modeling.

The elementary building block and “engine” of the continuous time stochastic models is the Brownian motion, named thus after the British botanist Robert Brown who observed the random movement of pollen particles in water.

Formally, a Wiener process is a random process $Z(t, w) : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ with increments that are statistically independent and normally distributed with mean zero and variance equal to the increment in time. In other words, for every pair of disjoint time intervals $[t_1, t_2], [t_3, t_4]$ with, say, $t_1 < t_2 \leq t_3 < t_4$, the increments $Z(t_4, w) - Z(t_3, w)$ and $Z(t_2, w) - Z(t_1, w)$ (also called white noise) are independent and normally distributed random variables with means:

$$E[Z(t_3, w) - Z(t_4, w)] = E[Z(t_2, w) - Z(t_1, w)] = 0 \quad (30.1)$$

MANAGING RISK IN THE MODERN WORLD

Applications of Bayesian Networks

By Norman Fenton and Martin Neil

Contents

	Page
Executive Summary	3
Bayesian Networks: Background and Development	4
What Can be Done with Bayesian Networks?	12
Current Challenges	16
Next Steps	21
Appendix 1: Active Researchers and Practitioner Groups	23
Appendix 2: BN tools	25
References	26

November 2007

A Knowledge Transfer Report from the London Mathematical Society and
the Knowledge Transfer Network for Industrial Mathematics

London Mathematical Society
De Morgan House, 57/58 Russell Square
London WC1B 4HS

Towards logistics systems parameter optimisation through the use of response surfaces

Katrien Ramaekers · Gerrit K. Janssens ·
Hendrik Van Landeghem

Received: 10 October 2005 / Revised: 16 January 2006 /
Published online: 29 September 2006
© Springer-Verlag 2006

Abstract Logistics systems have to cope with uncertainties in demand, in lead times, in transport times, in availability of resources and in quality. Management decisions have to take these uncertainties into consideration. An evaluation of decisions may be done by means of simulation. However, not all stochastic phenomena are of equal importance. By design of simulation experiments and making use of response surfaces, the most important phenomena are detected and their influence on performance estimated. Once the influence of the phenomena is known, this knowledge may be used to determine the optimal values of some decision parameters. An illustration is given on how to use response surfaces in a real-world case. A model is built in a logistics modelling software. The decision parameters have to be optimised for a specific objective function. Experiments are run to estimate the response surface. The validity of the response surface with few observations is also tested.

Keywords Simulation-Optimisation · Response surfaces · Experimental design · Regression · Logistics

K. Ramaekers (✉) · G. K. Janssens
Operations Management and Logistics, Hasselt University,
Agoralaan - Building D, 3590 Diepenbeek, Belgium
e-mail: katrien.ramaekers@uhasselt.be

G. K. Janssens
e-mail: gerrit.janssens@uhasselt.be

H. Van Landeghem
Industrial Management, Ghent University, Technologiepark 903,
9052 Zwijnaarde (Gent), Belgium
e-mail: hendrik.vanlandeghem@UGent.be

Title: Statistical modelling of financial time series: An introduction

Date: 8th March 2004

Year: 2004

Note no.: SAMBA/08/04

Author: Kjersti Aas <kjersti.aas@nr.no>
Xeni K. Dimakos <xeni.dimakos@nr.no>

Abstract: This note is intended as a summary of a one-day course in quantitative analysis of financial time series. It offers a guide to analysing and modelling financial time series using statistical methods, and is intended for researchers and practitioners in the finance industry.

Our aim is to provide academic answers to questions that are important for practitioners. The field of financial econometrics has exploded over the last decade. The intention of this course is to help practitioners cut through the vast literature on financial time series models, focusing on the most important and useful theoretical concepts.

Keywords: Statistical modelling, time series, stationarity, GARCH, correlation, copula

Target group: STAR

Availability: Open

Project: GB-BFF

Project no.: 220194

Research field: Finance, insurance and power market

No. of pages: 37

MIXED RESOLUTION DESIGNS AS ALTERNATIVES TO TAGUCHI INNER/OUTER ARRAY DESIGNS FOR ROBUST DESIGN PROBLEMS

C. M. BORROR AND D. C. MONTGOMERY*

Department of Industrial and Management Systems Engineering, Arizona State University, Tempe, AZ 85287-5906, USA

SUMMARY

There has been considerable debate over the contributions made by Genichi Taguchi to robust process and product design. As a result of the numerous debates, there have been many alternative approaches presented that are better suited to the robust design problem. In this paper a combined array design is presented as an alternative to a standard Taguchi design. The mixed resolution design is illustrated in an example involving control and noise variables. Two new variance properties of experimental designs are also presented. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: combined array design; mixed resolution design; robust parameter design; Taguchi methods

INTRODUCTION

Response surface methodology (RSM) lends itself to the design of products or processes that may be sensitive to uncontrollable or noise factors. For example, variables such as temperature and humidity may affect the performance of a process and are not necessarily controllable during routine process operation. It is important to consider the noise variables in the design or development stage of the process. By modelling the noise variables and control variables, a combination of settings can be determined for the control variables so that the process output will remain robust to changes in the noise variables.

Genichi Taguchi introduced the idea of robust parameter design [1–3]. He argued that not only should the controllable factors of interest x in a process be modelled, but also those uncontrollable or noise factors z that often cause variation in the response for a given experiment. Throughout the 1980s and 1990s there has been much debate over some of the design and analysis aspects of Taguchi's approach to robust design (see e.g. References [4–7]).

In his statistical design approach, Taguchi assigns the p control factors x and r noise factors z to an inner array and outer array respectively. Each array may be any number of standard designs known as orthogonal

arrays (OAs). The outer array consisting of, say, n_2 runs is then crossed with the inner array consisting of, say, n_1 runs, resulting in a total of $n_1 \times n_2$ total trials. The motivation for the crossed array strategy is straightforward. It allows all interactions between control variables and noise variables to be estimated. If such interactions exist, it should be possible in principle to find settings for the controllable variables that minimize the variability transmitted from the noise variables. In fact, without these interactions there will be no robust parameter design problem.

The experiment is carried out and the results are recorded and usually summarized in a response variable known as a signal-to-noise ratio (SNR). An analysis is performed to determine the best overall combination of levels of the control factors. Since the number of trials in this situation is potentially large, the control and noise arrays are usually small fractional factorial designs.

A typical analysis technique promoted by Taguchi includes the use of marginal means analysis and marginal means plots. One type of plot displays the SNRs plotted against the levels of each control factor. A second type of plot displays the average taken across all levels of the other factors again plotted against the level of the control factor.

There are significant drawbacks to the crossed array approach. For instance, along with the excessive number of runs that often results, the crossed array approach does not readily accommodate potentially

*Correspondence to: D. C. Montgomery, Department of Industrial and Management Systems Engineering, Arizona State University, Tempe, AZ 85287-5906, USA.

Preface

I developed this textbook while teaching the course *Statistics for Financial Engineering* to master's students in the financial engineering program at Cornell University. These students have already taken courses in portfolio management, fixed income securities, options, and stochastic calculus, so I concentrate on teaching statistics, data analysis, and the use of R, and I cover most sections of Chapters 4–9 and 17–20. These chapters alone are more than enough to fill a one semester course. I do not cover regression (Chapters 12–14 and 21) or the more advanced time series topics in Chapter 10, since these topics are covered in other courses. In the past, I have not covered cointegration (Chapter 15), but I will in the future. The master's students spend much of the third semester working on projects with investment banks or hedge funds. As a faculty adviser for several projects, I have seen the importance of cointegration.

A number of different courses might be based on this book. A two-semester sequence could cover most of the material. A one-semester course with more emphasis on finance would include Chapters 11 and 16 on portfolios and the CAPM and omit some of the chapters on statistics, for instance, Chapters 8, 18, and 20 on copulas, GARCH models, and Bayesian statistics. The book could be used for courses at both the master's and Ph.D. levels.

Readers familiar with my textbook *Statistics and Finance: An Introduction* may wonder how that volume differs from this book. This book is at a somewhat more advanced level and has much broader coverage of topics in statistics compared to the earlier book. As the title of this volume suggests, there is more emphasis on data analysis and this book is intended to be more than just “an introduction.” Chapters 8, 15, and 20 on copulas, cointegration, and Bayesian statistics are new. Except for some figures borrowed from *Statistics and Finance*, in this book R is used exclusively for computations, data analysis, and graphing, whereas the earlier book used SAS and MATLAB. Nearly all of the examples in this book use data sets that are available in R, so readers can reproduce the results. In Chapter 20 on Bayesian statistics, WinBUGS is used for Markov chain Monte Carlo and is called from R using

THE ANALYSIS OF TIME SERIES AN INTRODUCTION

Fifth edition

Chris Chatfield
Reader in Statistics
The University of Bath
United Kingdom

CHAPMAN & HALL/CRC

Boca Raton London New York Washington, D.C.

Disclaimer

In this book, I have tried to give an introductory overview of Monte Carlo methods in finance known to expert practitioners and, in places, I may not always have given due credit to all the pioneers who contributed to this borderline area of mathematics and finance. Wherever I fail to give explicit reference to the original inventor of any given method, this is not to mean I wish to pretend that it is my own development, it is merely my own laxness about the whole issue of referencing and citations. In fact, I may use phrases like 'I present below', etc. repeatedly, but they just stand for their literal meaning, namely that I present, not that I claim to have invented the particular method. I did consider it much more important to focus on an as-good-as-possible explanation of the techniques and mathematics, rather than spending time on exhaustive research through a whole string of references to establish who actually was the originator of the subject at hand. I include a rather-too-long bibliography at the end of the book, and I did try to reference and cite wherever I could see a direct link, but I may have failed many great researchers in the field of Monte Carlo methods by not referencing them in the right places, or not referencing them at all. *Mea culpa, mea maxima culpa.*

Econometrics Examination

1. Table 1 gives the quantity supplied of a commodity Y at various prices X , holding everything else constant. (a) Estimate the regression equation of Y on X . (b) Test for the statistical significance of the parameter estimates at the 5% level of significance. (c) Find R^2 and report all previous results in standard summary form. (d) Predict Y and calculate a 95% confidence or prediction interval for $X = 10$.

Table 1. Quantity Supplied at Various Prices

n	1	2	3	4	5	6	7	8
Y	12	14	10	13	17	12	11	15
X	5	11	7	8	11	7	6	9

2. Suppose that from 24 yearly observations on the quantity demand of a commodity in kilograms per year Y , its price in dollars X_1 , consumer's income in thousands of dollars X_2 , and the price of a substitute commodity in dollars X_3 , the following estimated regression is obtained, where the numbers in parentheses represent standard errors:

$$\hat{Y} = 13 - 7X_1 + 2.4X_2 - 4X_3$$

(2) (0.8) (18)

(a) Indicate whether the signs of the parameters conform to those predicted by demand theory. (b) Are the estimated slope parameters significant at the 5% level? (c) Find R^2 , if $\sum y^2 = 40$, $\sum yx_1 = 10$, and $\sum yx_2 = 45$ (where small letters indicate deviations from the mean). (d) Find \bar{R}^2 . (e) Is R^2 significantly different from zero at the 5% level? (f) Find the standard error of the regression. (g) Find the coefficient of price and income elasticity of demand at the means, given $\bar{Y} = 32$, $\bar{X}_1 = 8$, and $\bar{X}_2 = 16$.

3. When the level of business expenditures for new plants and equipment of nonmanufacturing firms in the United States Y_t from 1960 to 1979 is regressed on the GNP X_{1t} , and the consumer price index, X_{2t} , the following results are obtained:

$$\hat{Y}_t = 31.75 + 0.08 X_{1t} - 0.58 X_{2t} \quad R^2 = 0.98$$

(6.08) (-3.08) $d = 0.77$

(a) How do you know that autocorrelation is present? What is meant by *autocorrelation*? Why is autocorrelation a problem? (b) How can you estimate ρ , the coefficient of autocorrelation? (c) How can the value of ρ be used to transform the variables in order to correct for autocorrelation? How do you find the first value of the transformed variables? (d) Is there any evidence of remaining autocorrelation from the following results obtained by running the regression on the transformed variables (indicated by an asterisk)?

$$Y_t^* = 3.79 + 0.04 X_{1t}^* - 0.05 X_{2t}^* \quad R^2 = 0.96$$

(8.10) (-0.72) $d = 0.89$

What could be the cause of any remaining autocorrelation? How could this be corrected?

4. The following two equations represent a simple macroeconomic model:

$$R_t = a_0 + a_1 M_t + a_2 Y_t + u_{1t}$$

$$Y_t = b_0 + b_1 R_t + u_{2t}$$

where R is the interest rate, M is the money supply, and Y is income. (a) Why is this a simultaneous-equations model? Which are the endogenous and exogenous variables? Why would the estimation of the R and Y equations by OLS give biased and inconsistent parameter estimates? (b) Find the reduced form of the model. (c) Is this model underidentified, over-identified, or just identified? Why? What are the values of the structural coefficients? What

CHAPTER 10 —

Simultaneous-Equations Methods

10.1 SIMULTANEOUS-EQUATIONS MODELS

When the dependent variable in one equation is also an explanatory variable in some other equation, we have a *simultaneous-equations system or model*. The dependent variables in a system of simultaneous equations are called *endogenous variables*. The variables determined by factors outside the model are called *exogenous variables*. There is one *behavioral or structural equation* for each endogenous variable in the system (see Example 1). Using OLS to estimate the structural equations results in biased and inconsistent parameter estimates. This is referred to as *simultaneous-equations bias*. To obtain consistent parameter estimates, the *reduced-form equations* of the model must first be obtained. These express each endogenous variable in the system only as a function of the exogenous variable of the model (see Example 2).

EXAMPLE 1. The following two equations represent a simple macroeconomic model:

$$\begin{aligned} M_t &= a_0 + a_1 Y_t + u_{1t} \\ Y_t &= b_0 + b_1 M_t + b_2 I_t + u_{2t} \end{aligned}$$

where M_t is money supply in time period t , Y is income, and I is investment. Since M depends on Y in the first equation and Y depends on M (and I) in the second equation, M and Y are jointly determined, so we have a simultaneous-equations model. M and Y are the endogenous variables, while I is exogenous or determined outside the model. A change in u_{1t} affects M_t in the first equation. This, in turn, affects Y_t in the second equation. As a result, Y_t and u_{1t} are correlated, leading to biased and inconsistent OLS estimates of the M (and Y) equation.

EXAMPLE 2. The first reduced-form equation can be derived by substituting the second equation into the first and rearranging:

$$\begin{aligned} M_t &= a_0 + a_1(b_0 + b_1 M_t + b_2 I_t + u_{2t}) + u_{1t} \\ &= \frac{a_0 + a_1 b_0}{1 - a_1 b_1} + \frac{a_1 b_2}{1 - a_1 b_1} I_t + \frac{u_{1t} + a_1 u_{2t}}{1 - a_1 b_1} \end{aligned}$$

or

$$M_t = \pi_0 + \pi_1 I_t + v_{1t}$$

The second reduced-form equation can be derived by substituting the first equation into the second and rearranging:

Chapter 24

Applications of the Binomial Distribution to Evaluate Call Options

C Lee, John Lee, and Jessica Shin-Ying Mai

Abstract In this chapter, we first introduce the basic concepts of call and put options. Then we show how the simple one-period binomial call option pricing model can be derived. Finally, we show how a generalized binomial option pricing model can be derived.

Keywords Binomial distribution • Option • Simple binomial pricing model • Generalized binomial option pricing model • Hedge ratio • Cumulative binomial function • Exercise price • Decision tree

purchaser is under no obligation to buy; it is, indeed, an “option.” This attribute of an option contract distinguishes it from other financial contracts. For instance, whereas the holder of an option may let his or her claim expire unused if he or she so desires, other financial contracts (such as futures and forward contracts) obligate their parties to fulfill certain conditions.

A *call option* gives its owner the right to buy the underlying security, a *put option* the right to sell. The price at which the stock can be bought (for a call option) or sold (for a put option) is known as the exercise price.

24.1 Introduction

In this chapter, we will show that how the binomial distribution can be used to derive the call option pricing model. In the second section of this chapter, we will define the basic concept of option. In the third section, we will define and derive the simple binomial option pricing model. In the fourth section, we will derive the generalized n period binomial option pricing model. Finally, in the fifth section, we will summarize our findings and make concluding remarks.

24.2 What Is an Option?

In the most basic sense, an **option** is a contract conveying the right to buy or sell a designated security at a stipulated price. The contract normally expires at a predetermined date. The most important aspect of an option contract is that the

24.3 The Simple Binomial Option Pricing Model

Before discussing the binomial option model, we must recognize its two major underlying assumptions. First, the binomial approach assumes that trading takes place in discrete time; that is, on a period-by-period basis. Second, it is assumed that the stock price (the price of the underlying asset) can take on only two possible values each period; it can go up or go down.

Say we have a stock whose current price per share S can advance or decline during the next period by a factor of either u (up) or d (down). This price either will increase by the proportion $u - 1 \geq 0$ or will decrease by the proportion $1 - d$, $0 < d < 1$. Therefore, the value S in the next period will be either uS or dS . Next, suppose that a call option exists on this stock with a current price per share of C and an exercise price per share of X and that the option has one period left to maturity. This option’s value at expiration is determined by the price of its underlying stock and the exercise price X . the value is either

$$C_u = \max(0, uS - X) \quad (24.1)$$

Or

$$C_d = \max(0, dS - X) \quad (24.2)$$

Rutgers Corp., Boston, MA, USA

for PBBEF Research, New York, NY, USA

J. Mai (✉)

Rutgers University, Newark, NJ, USA

jmai@eden.rutgers.edu

Chapter 16

Portfolio Analysis

Jack Clark Francis

Abstract In 1952, Harry M. Markowitz published a seminal paper about analyzing portfolios. In 1990, he was awarded the Nobel Prize for his portfolio theory. Markowitz portfolio analysis delineates a set of highly desirable investment portfolios. These optimal portfolios have the maximum return at each plausible level of risk, computed iteratively over a range of different risk levels. Conversely, Markowitz portfolio analysis can find the same set of optimal investments by delineating portfolios that have the minimum risk over a range of different rates of return. The set of all Markowitz optimal portfolios is called the efficient frontier. Portfolio analysis analyzes rate of return statistics, risk statistics (standard deviations), and correlations from a list of candidate investments (stocks, bonds, and so on) to determine which investments, and in what proportions (weights), enter into every efficient portfolio. Further analysis of Markowitz's portfolio theory reveals interesting asset pricing implications.

Keywords Portfolio • Rate of return • Portfolio analysis • Standard deviation • Correlation • Weight • Variance-covariance matrix • Risk minimization • Return maximization • Efficient portfolios • Efficient frontier

16.1 Introduction

Portfolio analysis is a mathematical procedure that determines a selection of optimum portfolios for investors to consider. The procedure was first made public in 1952 by Harry Markowitz.¹ The theory caused a scientific revolution in finance. Before Markowitz's scientific procedure investment counselors passed off "common sense guidelines" to their clients as valuable expert advice.

J.C. Francis (✉)
Baruch College, New York, NY, USA
e-mail: Jack@JCFrancis.com

¹ The analysis was originally presented in an article: Markowitz (1952). Markowitz expanded his presentation in a book, Markowitz (1959).

The objective of portfolio analysis is to determine a set of "efficient portfolios." Figure 16.1 is a two-dimensional graph in risk-return space that depicts the set of efficient portfolios as a curve between points E to F. This curve is called the efficient frontier; it is a set of optimum investment portfolios generated by analyzing an underlying group of individual assets. The assets that are being analyzed can be thought of as being individual issues of stocks and bonds. These individual assets are represented by the dots lying below the efficient frontier in Fig. 16.1.

16.2 Inputs for Portfolio Analysis

Portfolio analysis requires certain data as inputs. If n individual assets (stocks and bonds) are the potential investments, the inputs to portfolio analysis are as follows:

1. n expected returns.
2. n variances of returns.
3. $(n^2 - n)/2$ covariances, or correlations.

For instance, to generate a three-security portfolio, the analysis requires the statistics listed in Table 16.1.

16.3 The Security Analyst's Job

The security analyst needs to analyze "one period" rates of return defined in Equation (16.1). For example, consider James's purchase of a stock for \$54 per share. He sold that stock 1 year later for \$64 to realize a capital gain of \$10 and a cash dividend of 80 cents per share. Jim's total income from the stock is \$10.80 and his rate of return is 20%.

Chapter 10

Portfolio Optimization Models and Mean–Variance Spanning Tests

Wei-Peng Chen, Huimin Chung, Keng-Yu Ho, and Tsui-Ling Hsu

Abstract In this chapter we introduce the theory and the application of the computer program of modern portfolio theory. The notion of diversification is age-old: “don’t put your eggs in one basket,” obviously predates economic theory. However, a formal model showing how to make the most of the power of diversification was not devised until 1952, a feat for which Harry Markowitz eventually won the Nobel Prize in economics. Markowitz portfolio shows that as you add assets to an investment portfolio the total risk of that portfolio – as measured by the variance (or standard deviation) of total return – declines continuously, but the expected return of the portfolio is a weighted average of the expected returns of the individual assets. In other words, by investing in portfolios rather than in individual assets, investors could lower the total risk of investing without sacrificing return. In the second part we introduce the mean–variance spanning test that follows directly from the portfolio optimization problem.

Keywords Minimum variance portfolio • Optimal risky portfolio • Capital allocation line • Mean–variance spanning tests

10.1 Introduction of Markowitz Portfolio-Selection Model

Harry Markowitz (1952, 1959) developed his portfolio-selection technique, which came to be called *modern portfolio theory* (MPT). Prior to Markowitz’s work, security-selection models focused primarily on the returns generated

W.-P. Chen (✉)

Department of Finance, Shih Hsin University, Taipei, Taiwan
e-mail: wpchen@cc.shu.edu.tw

H. Chung and T.-L. Hsu

Graduate Institute of Finance, National Chiao Tung University,
Hsinchu, Taiwan
e-mail: chunhui@mail.nctu.edu.tw; tracy.shu@msa.hinet.net

K.-Y. Ho

Department of Finance, National Taiwan University, Taipei, Taiwan
e-mail: kengyuho@management.ntu.edu.tw

by investment opportunities. Standard investment advice was to identify those securities that offered the best opportunity for gain with the least risk and then construct a portfolio from these. Following this advice, an investor might conclude that railroad stocks all offered good risk-reward characteristics and compile a portfolio entirely from these. The Markowitz theory retained the emphasis on return; but it elevated risk to a coequal level of importance, and the concept of portfolio risk was born. Whereas risk has been considered an important factor and variance an accepted way of measuring risk, Markowitz was the first to clearly and rigorously show how the variance of a portfolio can be reduced through the impact of diversification. He proposed that investors focus on selecting portfolios based on their overall risk-reward characteristics instead of merely compiling portfolios from securities that each individually have attractive risk-reward characteristics.

A *Markowitz portfolio model* is one where no added diversification can lower the portfolio’s risk for a given return expectation (alternately, no additional expected return can be gained without increasing the risk of the portfolio). The *Markowitz Efficient Frontier* is the set of all portfolios for which expected returns reach the maximum given a certain level of risk.

The *Markowitz model* is based on several assumptions concerning the behavior of investors and financial markets:

1. A probability distribution of possible returns over some holding period can be estimated by investors.
2. Investors have single-period utility functions in which they maximize utility within the framework of diminishing marginal utility of wealth.
3. Variability about the possible values of return is used by investors to measure risk.
4. Investors care only about the means and variance of the returns of their portfolios over a particular period.
5. Expected return and risk as used by investors are measured by the first two moments of the probability distribution of returns—expected value and variance.

Chapter 5

Risk-Aversion, Capital Asset Allocation, and Markowitz Portfolio-Selection Model

Cheng-Few Lee, Joseph E. Finnerty, and Hong-Yi Chen

Abstract In this chapter, we first introduce utility function and indifference curve. Based on utility theory, we derive the Markowitz's model and the efficient frontier through the creation of efficient portfolios of varying risk and return. We also include methods of solving for the efficient frontier both graphically and mathematically, with and without explicitly incorporating short selling.

Keywords Markowitz model • Utility theory • Utility functions • Indifference curve • Risk averse • Short selling • Dyl model • Iso-return line • Iso-variance ellipse • Critical line • Lagrange multipliers

5.1 Introduction

In this chapter, we address basic portfolio analysis concepts and techniques discussed in the Markowitz portfolio-selection model and other related issues in portfolio analysis. Before Harry Markowitz (1952, 1959) developed his portfolio-selection technique into what is now modern portfolio theory (MPT), security-selection models focused primarily on the returns generated by investment opportunities. The Markowitz theory retained the emphasis on return, but it elevated risk to a coequal level of importance, and the concept of portfolio risk was born. Whereas risk had been considered an important factor and variance an accepted way of measuring risk, Markowitz was the first to clearly and rigorously show how the variance of a portfolio can be reduced through the impact of diversification. He demonstrated that by combining securities that are not perfectly positively correlated into a portfolio, the portfolio variance can be reduced.

The *Markowitz model* is based on several assumptions concerning the behavior of investors:

H.-Y. Chen (✉) and C.-F. Lee
Rutgers University, New Brunswick, NJ, USA
e-mail: hchen37@pegasus.rutgers.edu

J.E. Finnerty
University of Illinois, Urbana-Champaign, IL, USA
e-mail: finnerty@illinois.edu

1. A probability distribution of possible returns over some holding period can be estimated by investors.
2. Investors have single-period utility functions in which they maximize utility within the framework of diminishing marginal utility of wealth.
3. Variability about the possible values of return is used by investors to measure risk.
4. Investors use only expected return and risk to make investment decisions.
5. Expected return and risk as used by investors are measured by the first two moments of the probability distribution of returns—expected value and variance.
6. Return is desirable; risk is to be avoided.

It follows, then, that a security or portfolio is considered efficient if there is no other investment opportunity with a higher level of return at a given level of risk and no other opportunity with a lower level of risk at a given level of return.

5.2 Measurement of Return and Risk

This section focuses on the return and risk measurements utilized in applying the Markowitz model to efficient portfolio selection.

5.2.1 Return

Using the probability distribution of expected returns for a portfolio, investors are assumed to measure the level of return by computing the *expected value* of the distribution.

$$E(R_P) = \sum_{i=1}^n W_i E(R_i) \quad (5.1)$$

where:

$$\sum_{i=1}^n W_i = 1.0;$$

Chapter 1

Theoretical Framework of Finance

Abstract The main purpose of this chapter is to explore important finance theories. First, we discuss discounted cash-flow valuation theory (classical financial theory). Second, we discuss the Modigliani and Miller (M and M) valuation theory. Third, we examine Markowitz portfolio theory. We then move on to the capital asset pricing model (CAPM), followed by the arbitrage pricing theory. Finally, we will look at the option pricing theory and futures valuation and hedging.

Keywords Discounted cash-flow valuation • M and M valuation theory • Markowitz portfolio theory • Capital asset pricing model • Arbitrage pricing theory • Option pricing model • Futures valuation and hedging

1.1 Introduction

Value determination of financial instruments is important in security analysis and portfolio management. Valuation theories are the basic tools for determining the intrinsic value of alternative financial instruments. This chapter provides a general review of the financial theory that most students of finance would have already received in basic corporate finance and investment classes. Synthesis and integration of the valuation theories are necessary for the student of investments in order to have a proper perspective of security analysis and portfolio management.

The basic policy areas involved in the management of a company are (1) investment policy, (2) financial policy, (3) dividend policy, and (4) production policy. Since the determination of the market value of a firm is affected by the way management sets and implements these policies, they are of critical importance to the security analyst. The security analyst must evaluate management decisions in each of these areas and convert information about company policy into price estimates of the firm's securities. This chapter examines these policies within a financial theory framework, dealing with valuation models.

There are six alternative but interrelated valuation models of financial theory that might be useful for the analysis of securities and the management of portfolios:

1. Discounted cash-flow valuation theory (classical financial theory)
2. M and M valuation theory
3. Capital asset pricing model (CAPM)
4. Arbitrage Pricing Theory (APT)
5. Option-pricing theory (OPT)
6. Futures Valuation and Hedging

The discounted cash-flow valuation and M and M theories are discussed in the typical required corporate-finance survey course for both bachelor's and master's programs in business. The main purpose of this chapter is to review these theories and discuss their interrelationships. The discounted cash-flow model is first reviewed by some of the basic valuation concepts in Sect. 1.2. In the second section, the four alternative evaluation methods developed by M and M in their 1961 article are discussed. Their three propositions and their revision with taxes are explored, including possible applications of their theories in security analysis. Miller's inclusion of personal taxes is discussed in Sect. 1.3. Section 1.4 discusses the Markowitz portfolio theory. Section 1.5 includes a brief overview of CAPM concepts. Section 1.6 introduces the Arbitrage Pricing Theory (APT). Sections 1.6 and 1.7 discuss the option-pricing theory and the futures valuation and hedging. Conclusion is presented in Sect. 1.8.

1.2 Discounted Cash-Flow Valuation Theory

Discounted cash-flow valuation theory is the basic tool for determining the theoretical price of a corporate security. The price of a corporate security is equal to the present value of future benefits of ownership. For example, for common stock, these benefits include dividends received while the stock is owned plus capital gains earned during the ownership period. If we assume a one-period investment and a world of certain cash flows, the price paid for a share of