
Half Title Page

Title Page

LOC Page

*To my dog
and my cat.*

Contents

| | |
|---|-----------|
| Foreword | xi |
| Preface | xiii |
| Contributors | xv |
| SECTION I This is a Part | |
| CHAPTER 11 ■ Introduction to Image Processing and the MATLAB Environment | 3 |
| 11.1 INTRODUCTION | 4 |
| 11.1.1 A Component Part | 4 |
| 11.2 RECORD LINKAGE MODEL | 8 |
| 11.2.1 A Component Part | 8 |
| 11.2.1.1 H3 A Component Part | 8 |
| CHAPTER 12 ■ Introduction to Image Processing and the MATLAB Environment | 17 |
| 12.1 INTRODUCTION | 18 |
| 12.1.1 A Component Part | 18 |
| 12.2 RECORD LINKAGE MODEL | 22 |
| 12.2.1 A Component Part | 22 |
| 12.2.1.1 H3 A Component Part | 22 |
| CHAPTER 13 ■ Introduction to Image Processing and the MATLAB Environment | 31 |
| 13.1 INTRODUCTION | 32 |
| 13.1.1 A Component Part | 32 |
| 13.2 RECORD LINKAGE MODEL | 36 |

| | |
|---|-----------|
| 13.2.1 A Component Part | 36 |
| 13.2.1.1 H3 A Component Part | 36 |
| CHAPTER 14 ■ Introduction to Image Processing and the MATLAB Environment | 45 |
| <hr/> | |
| 14.1 INTRODUCTION | 46 |
| 14.1.1 A Component Part | 46 |
| 14.2 RECORD LINKAGE MODEL | 50 |
| 14.2.1 A Component Part | 50 |
| 14.2.1.1 H3 A Component Part | 50 |
| CHAPTER 15 ■ Introduction to Image Processing and the MATLAB Environment | 59 |
| <hr/> | |
| 15.1 INTRODUCTION | 60 |
| 15.1.1 A Component Part | 60 |
| 15.2 RECORD LINKAGE MODEL | 64 |
| 15.2.1 A Component Part | 64 |
| 15.2.1.1 H3 A Component Part | 64 |
| CHAPTER 16 ■ Introduction to Image Processing and the MATLAB Environment | 73 |
| <hr/> | |
| 16.1 INTRODUCTION | 74 |
| 16.1.1 A Component Part | 74 |
| 16.2 RECORD LINKAGE MODEL | 78 |
| 16.2.1 A Component Part | 78 |
| 16.2.1.1 H3 A Component Part | 78 |
| CHAPTER 17 ■ Introduction to Image Processing and the MATLAB Environment | 87 |
| <hr/> | |
| 17.1 INTRODUCTION | 88 |
| 17.1.1 A Component Part | 88 |
| 17.2 RECORD LINKAGE MODEL | 92 |
| 17.2.1 A Component Part | 92 |
| 17.2.1.1 H3 A Component Part | 92 |

| | |
|---|------------|
| CHAPTER 18 ■ Introduction to Image Processing and the MATLAB Environment | 101 |
| 18.1 INTRODUCTION | 102 |
| 18.1.1 A Component Part | 102 |
| 18.2 RECORD LINKAGE MODEL | 106 |
| 18.2.1 A Component Part | 106 |
| 18.2.1.1 H3 A Component Part | 106 |
| CHAPTER 19 ■ Introduction to Image Processing and the MATLAB Environment | 115 |
| 19.1 INTRODUCTION | 116 |
| 19.1.1 A Component Part | 116 |
| 19.2 RECORD LINKAGE MODEL | 120 |
| 19.2.1 A Component Part | 120 |
| 19.2.1.1 H3 A Component Part | 120 |
| CHAPTER 20 ■ Introduction to Image Processing and the MATLAB Environment | 129 |
| 20.1 INTRODUCTION | 130 |
| 20.1.1 A Component Part | 130 |
| 20.2 RECORD LINKAGE MODEL | 134 |
| 20.2.1 A Component Part | 134 |
| 20.2.1.1 H3 A Component Part | 134 |
| CHAPTER 21 ■ Introduction to Image Processing and the MATLAB Environment | 143 |
| 21.1 INTRODUCTION | 144 |
| 21.1.1 A Component Part | 144 |
| 21.2 RECORD LINKAGE MODEL | 148 |
| 21.2.1 A Component Part | 148 |
| 21.2.1.1 H3 A Component Part | 148 |
| Bibliography | 157 |
| Index | 159 |

Foreword

Abóritiorior maio. Optaess inullorē cum aítate niate nonem ut ario. Ut labor soloreped magniae ipsum etur sunt magnatiopos nonsequi ut aliquideles quam quisque magnatia voluptae poreperiat. Ibus, num escilla boreribus, uteum qui arum fugitae rrorit que ide volum debit faci conseque nus as quo doluptatur autem qui odis dero esequis nihillaborio di audit ut moluptate opta dolorere vellabore verum fugita del molori occus pores est et fugiati audis enimus sitentur? Possequ asperupita sedit, int. Im dolo vid quo te quis dollam re ne que sime parchit aspici unto te accusam eium eium reptas sitatio etur? Gent pereperum endunt magnate pa dolore cuptatent dolum aliquae eturera voluptat vent odiatinusaes quos aliquias minvel ipsunt adicid eossunto tet es ex estios sit volupta commo tendemp oribus, aut in nimped quam utate suntiam enis eictatem doluptatem volorendis archilibus. Ucipsan dissimp eriatio dis qui aspe ium et anda consequis min rest, vit ad que liqui asitatur, nonsed moles dolorum illacienis adit quias in eiciet doluptasint doluptisque volupta senihillaut amus ut et es undic tem velit estrunt et pra quaerchit ut opti rero ipis modio tendae prorat ea velit ut es es corruptatur molent explabo ruptat labo. Et aut quos que dolut velibuscil maxim debite exces moluptatium volupiendam nullignatus. Qui omnit pla ium qui alicim repe vendi delenim intur, quia volorum, esti te pa cusam que simi, tem et, eosam, conecae nim vent. Ficidebit fugitatis reptat lant. Mos voluptas sit officiminci bea et inci que estrunt maximet a peliqui aspeliquos et aperior re saperum rherchia quideri volorro quia veliquam faceris ma voloreptati qui omnis mod mo tem repediciis rem. Et quias molorit imillam ne pra ipsam id moles min nonsequos exerum volest et, temporepudae sunt rem nobit modi dolorepta volenim dolupti beatiberum ipsuntenis ditenis evellaut omnima volorit eosamet fugiati volorro blaborem quis accessunt.

Um con commolupta si conecto rporum dolliqui ut odit ut prem. Inhibit erorior eptur, sitaten esequi atiaestis que vel ium quosae prerunt issimus mi, aborat ium voluptasin nulpa quam lique pro dus imil est liaturi onseque num am, site num eactecto taquid magnim alitiun diciaspit volor aspe nimagnis anis autati quoditis que volupta tiant, eatio. Et utet et volupta turiorum facillabo. Elique pore etur sit explant atibusciet adit venimuscum voluptibus pro et et estia venditiumqui nonessitist, sim ad qui int, que eatiorent ipsant landior maximet, tempore mpedis aciuntio tem sunt reped etur? Ugit, aut la et ese pligendebis si quam quam nonsequ atquodio con nem porrero mil illestiatet apic tet qui conectia vollab ipiciis et il es dolupta quaspedis quo et rerunt inum et ea delenih ilistempero beri culpa doluptur, el maximenimus modias as di del ima con repuda vita eat iduntis magnatent enda voluptassin re es molorecerum quissit lab ipsanda eptate samus et qui utatem eatium faces adictur?

Ficate et et quam laccatinum et a sinverepe core cum quat. Fugiam name com mos volupta tiusantus voluptatium rerum reius consendi des aut la doluptate si que

sunt aut odi commnia parunt voluptatur, aut ressitiore et aut liberit atibusdae cuptati nctotas dendebit et isquaspiento doluptatia nonet ad quo omnisit ibusam, valorib usamus. Da quam fugiatem volupta quaesequodi ni tem quia vit idundende est, imo quis que am lam nos molorep ellenienis dolorecusdae re nossus autenditi derchicus nates sum haritate liaturit, ulliquam cum, susa corepudam harum eaquodicimin corenducit, sit, coreperro moloreperem eventem. Ignation nonsequias moluptam quias modia et, tem. Ut erupti doluptatem volorro corerum quae plabore mprrorum fugias arum laborionsed maximod et hilia autas aspis aut ipsum rest earchilis enecero tem res desere con restemporem faceped ma dolupta sperore mperectem. Nempelenim nullore corest, quatus aut maximus quunt illaboremped quidio voluptatur?

Ati volorepudist laborrum nam, commolecto il ius. Eni omnit labor maxim dolores as as sit as et libusapis veliqui busiae quam estius, contemporibus ant ut as consedi gendaerum volum non nossunt eum doles im quibus de pe non rem vendit, sunt autem eiusam dolupta dia dellesequi iunders pellamus. Cones quias que oditaec tessitaepre rem ent es moluptatius. Aximinus ipsae si alibus. Onsed magnimuscit rercid maionsed que nosaerum re et ellucti busant anim quide iur atust as atqui offic tet ut odicimus ea pratia deribustio. Tection sernatis dolo magnat lamusdae sunt. Otat labo. Ut que sit latque ped moluptio ea doloressed eatusaperiam eosse nonsequi aces suntur rempore mporepe dition preheni magnihi ligendant foccus.

Preface

Aboritior maio. Optaess inullore cum autate niate nonem ut ario. Ut labor soloreped magniae ipsum etur sunt magnatiopos nonsequi ut aliquideles quam quisque magnatia voluptae poreperiat. Ibus, num escilla boreribus, utem qui arum fugitae rrorit que ide volum debit faci conseque nus as quo doluptatur autem qui odis dero esequis nihillaborio di audit ut moluptate opta dolorere vellabore verum fugita del molori occus pores est et fugiati audis enimus sitentur? Possequ asperupita sedit, int. Im dolo vid quo te quis dollam re ne que sime parchit aspici unto te accusam eium eium reptas sitatio etur? Gent pereperum endunt magnate pa dolore cuptatent dolum aliquae eturera voluptat vent odiatinusaes quos aliquias minvel ipsunt adicid eossumt tet es ex estios sit volupta commo tendemp oribus, aut in nimped quam utate suntiam enis eictatem doluptatem volorendis archilibus. Ucipsan dissimp eriatio dis qui aspe ium et anda consequis min rest, vit ad que liqui asitatur, nonsed moles dolorum illaciensis adit quias in eiciet doluptasint doluptisque volupta senihillaut amus ut et es undic tem velit estrunt et pra quaerchit ut opti rero ipis modio tendae prorat ea velit ut es es corruptatur molent explabo ruptat labo. Et aut quos que dolut velibuscil maxim debite exces moluptatium volupiendam nullignatus. Qui omnit pla ium qui alicim repe vendi delenim intur, quia valorum, esti te pa cusam que simi, tem et, eosam, conecae nim vent. Ficidebit fugitatis reptat lant. Mos voluptas sit officiminci bea et inci que estrunt maximet a peliqui aspeliquos et aperior re saperum rherchia quideri valorro quia veliquam faceris ma voloreptati qui omnis mod mo tem repediciis rem. Et quias molorit imillam ne pra ipsam id moles min nonsequos exerum volest et, temporepudae sunt rem nobit modi dolorepta volenim dolupti beatiberum ipsuntenis ditenis evellaut omnima valorit eosamet fugiati valorro blaborem quis accessunt.

Um con commolupta si conecto rporum dolliqui ut odit ut prem. Inhibit erorior eptur, sitaten esequi atiaestis que vel ium quosae prerunt issimus mi, aborat ium voluptasin nulpa quam lique pro dus imil est liaturi onseque num am, site num eactecto taqid magnim alitium diciaspit valor aspe nimagnis anis autati quoditis que volupta tiant, eatio. Et utet et volupta turiorum facillabo. Elique pore etur sit explant atibusciet adit venimuscum voluptibus pro et et estia venditiumqui nonessist, sim ad qui int, que eatiorent ipsant landior maximet, tempore mpedis aciuntio tem sunt reped etur? Ugit, aut la et ese pligendebis si quam quam nonsequ atquodio con nempororero mil illestiatet apic tet qui conectia vollab ipiciis et il es dolupta quaspedis quo et rerunt inum et ea delenih ilistempero beri culpa doluptur, el maximenimus modias as di del ima con repuda vita eat iduntis magnatent enda voluptassin re es molorecerum quissit lab ipsanda eptate samus et qui utatem eatium faces adictur?

Ficate et et quam laccatinum et a sinverepe core cum quat. Fugiam name commos volupta tiusantus voluptatium rerum reius consendi des aut la doluptate si que

sunt aut odi commnia parunt voluptatur, aut ressitiore et aut liberit atibusdae cuptati nctotas dendebit et isquaspiento doluptatia nonet ad quo omnisit ibusam, valorib usamus. Da quam fugiatem volupta quaesequodi ni tem quia vit idundende est, imo quis que am lam nos molorep ellenienis dolorecusdae re nossus autenditi derchius nates sum haritate liaturit, ulliquam cum, susa corepudam harum eaquodicimin corenducit, sit, coreperro moloreperem eventem. Ignation nonsequias moluptam quias modia et, tem. Ut erupti doluptatem volorro corerum quae plabore mporrum fugias arum laborionsed maximod et hilia autas aspis aut ipsum rest earchilis enecero tem res desere con restemporem faceped ma dolupta sperore mperectem. Nempelenim nullore corest, quatus aut maximus quunt illaboremped quidio voluptatur?

Ati volorepudist laborrum nam, commolecto il ius. Eni omnit labor maxim dolores as as sit as et libusapis veliqui busiae quam estius, contemporibus ant ut as consedi gendaerum volum non nossunt eum doles im quibus de pe non rem vendit, sunt autem eiusam dolupta dia dellesequi iunders pellamus. Cones quias que oditaec tessitaepre rem ent es moluptatius. Aximinus ipsae si alibus. Onsed magnimuscit rercid maionsed que nosaerum re et ellucti busant anim quide iur atust as atqui offic tet ut odicimus ea pratia deribustio. Tection sernatis dolo magnat lamusdae sunt. Otat labo. Ut que sit latque ped moluptio ea doloressed eatusaperiam eosse nonsequi aces suntur rempore mporepe dition preheni magnihi ligendant foccus.

Contributors

Michael Aftosmis

NASA Ames Research Center
Moffett Field, California

Pratul K. Agarwal

Oak Ridge National Laboratory
Oak Ridge, Tennessee

Sadaf R. Alam

Oak Ridge National Laboratory
Oak Ridge, Tennessee

Gabrielle Allen

Louisiana State University
Baton Rouge, Louisiana

Martin Sandve Aln_s

Simula Research Laboratory and University
of Oslo, Norway
Norway

Steven F. Ashby

Lawrence Livermore National Laboratory
Livermore, California

David A. Bader

Georgia Institute of Technology
Atlanta, Georgia

Benjamin Bergen

Los Alamos National Laboratory
Los Alamos, New Mexico

Jonathan W. Berry

Sandia National Laboratories
Albuquerque, New Mexico

Martin Berzins

University of Utah

Salt Lake City, Utah

Abhinav Bhatele

University of Illinois
Urbana-Champaign, Illinois

Christian Bischof

RWTH Aachen University
Germany

Rupak Biswas

NASA Ames Research Center
Moffett Field, California

Eric Bohm

University of Illinois
Urbana-Champaign, Illinois

James Bordner

University of California, San Diego
San Diego, California

George Bosilca

University of Tennessee
Knoxville, Tennessee

Greg L. Bryan

Columbia University
New York, New York

Marian Bubak

AGH University of Science and Technology
Kraków, Poland

Andrew Canning

Lawrence Berkeley National Laboratory
Berkeley, California

Jonathan Carter

Lawrence Berkeley National Laboratory
Berkeley, California

Piscataway, New Jersey

Zizhong Chen

Jacksonville State University
Jacksonville, Alabama

Roxana E. Diaconescu

Yahoo! Inc.
Burbank, California

Joseph R. Crobak

Rutgers, The State University of New
Jersey

Roxana E. Diaconescu

Yahoo! Inc.
Burbank, California

I

This is a Part

Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \{\epsilon\}$
- $\mathbb{E} = \{\epsilon\}$
- $\emptyset = \{\circ\}$

11.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

11.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 11.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure *k*-anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table *PT* and a set of tables T_1, \dots, T_n , where *PT* contains a person identifier *Pid* and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in *QID*, and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same *QID* with o . It is important to emphasize that the *k*-anonymization is applied at the *record owner* level, not at the *record* level in traditional *k*-anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 11.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \widehat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\widehat{\alpha}_j - \widehat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (11.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (11.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 11.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 11.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

11.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

11.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

11.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

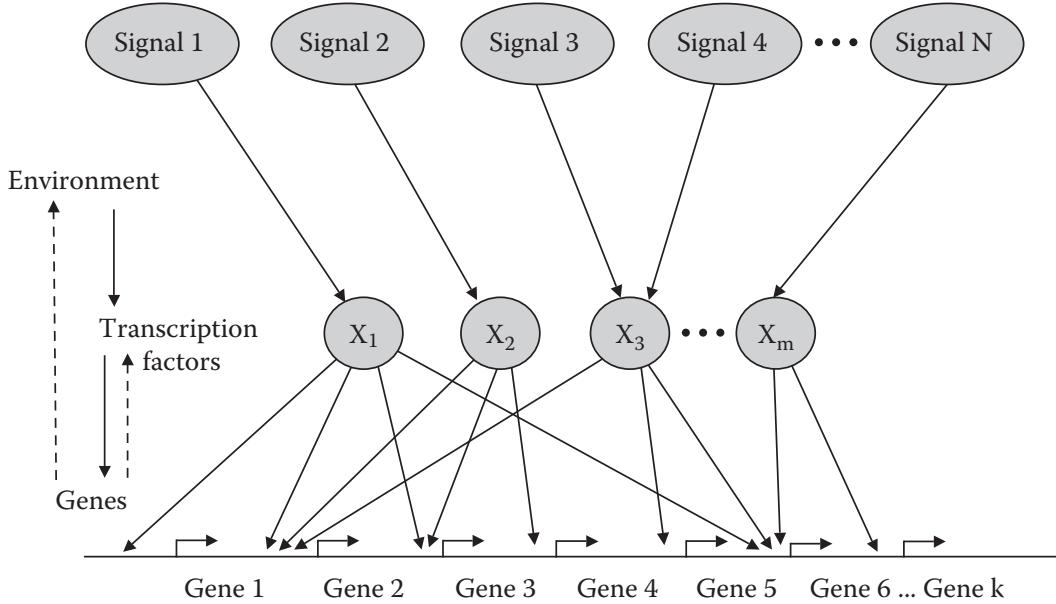


Figure 11.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 11.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

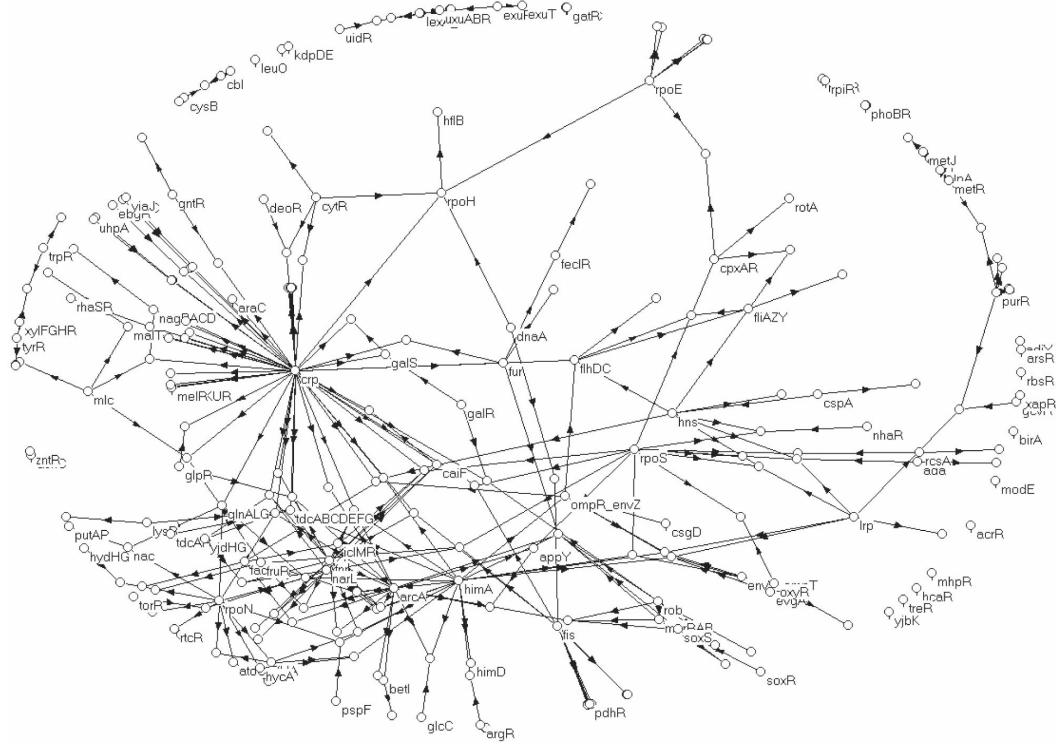


Figure 11.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

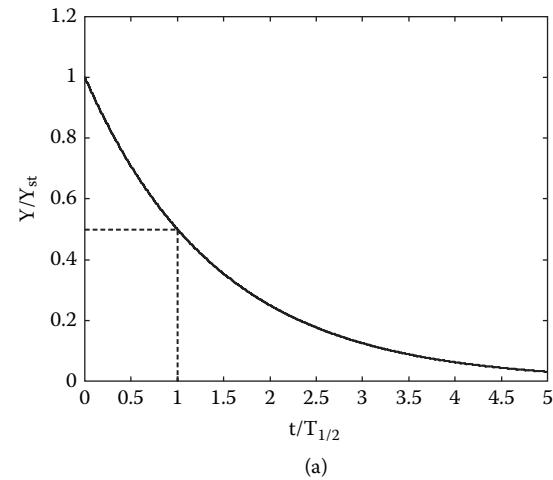
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

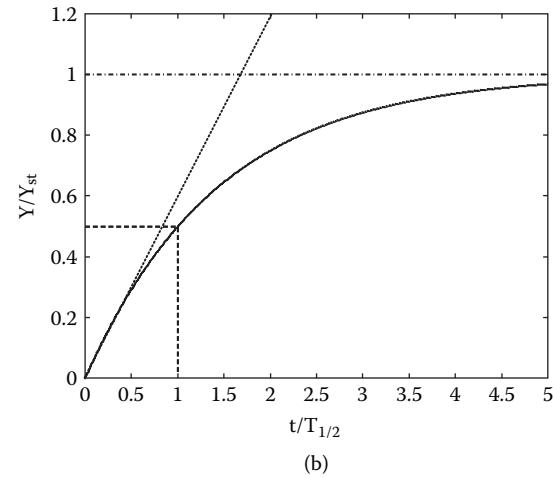
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 11.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 11.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 11.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P11.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 11.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P11.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P11.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P11.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P11.5})$$

- 11.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

12.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

12.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 12.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 12.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (12.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (12.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 12.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 12.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

12.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

12.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

12.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

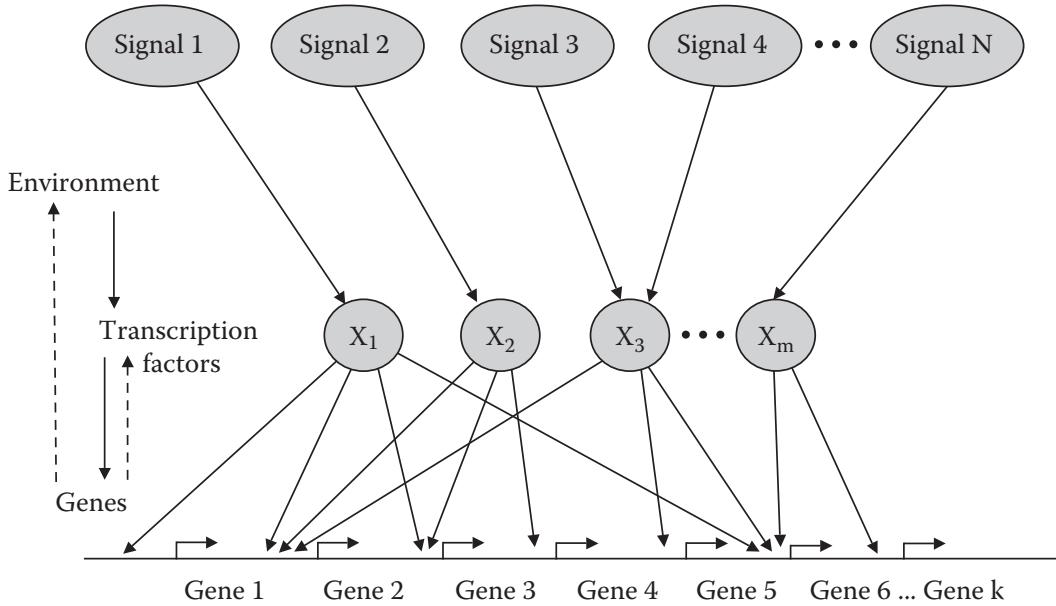


Figure 12.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 12.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

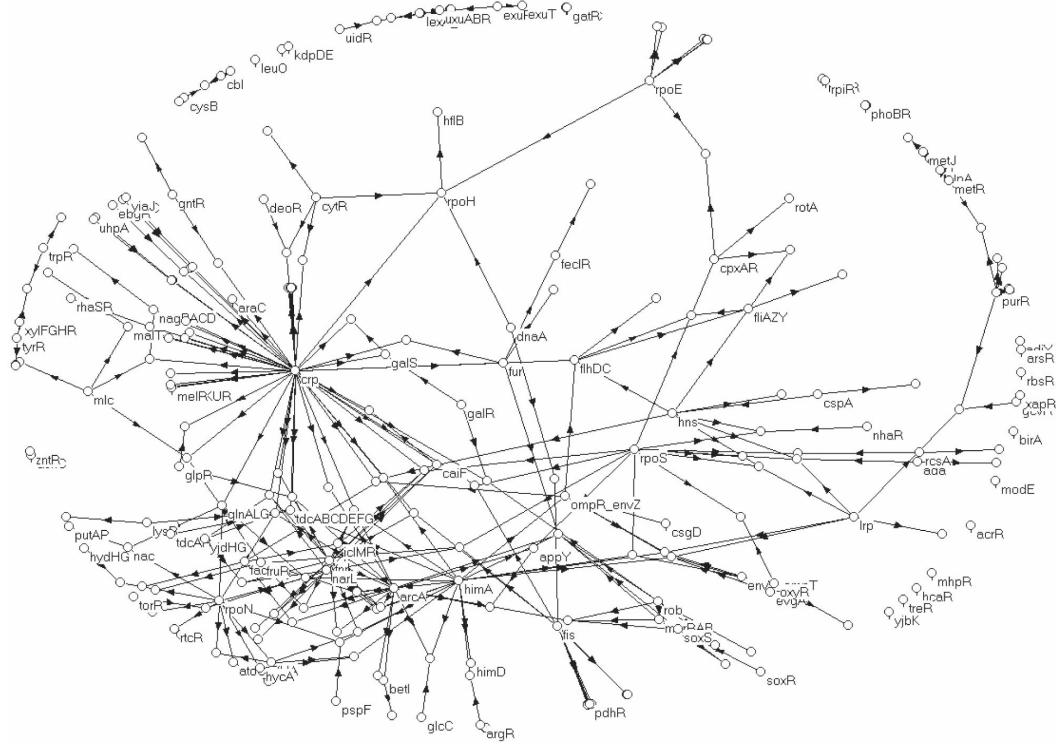


Figure 12.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

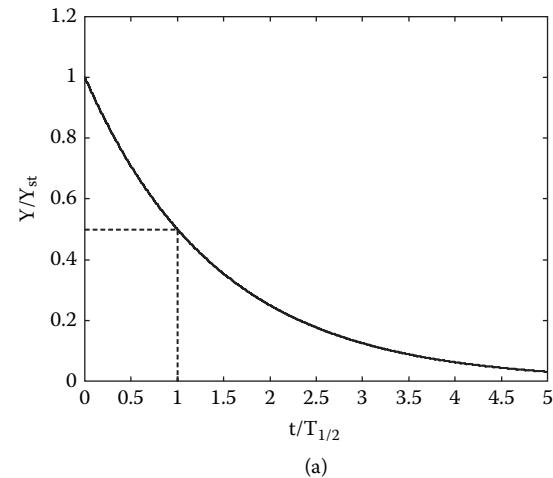
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

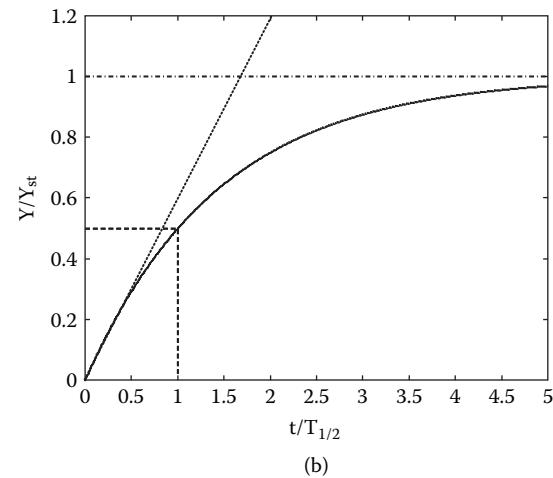
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 12.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 12.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 12.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P12.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 12.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P12.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P12.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P12.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P12.5})$$

- 12.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

13.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

13.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 13.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 13.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (13.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (13.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 13.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 13.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

13.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

13.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

13.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

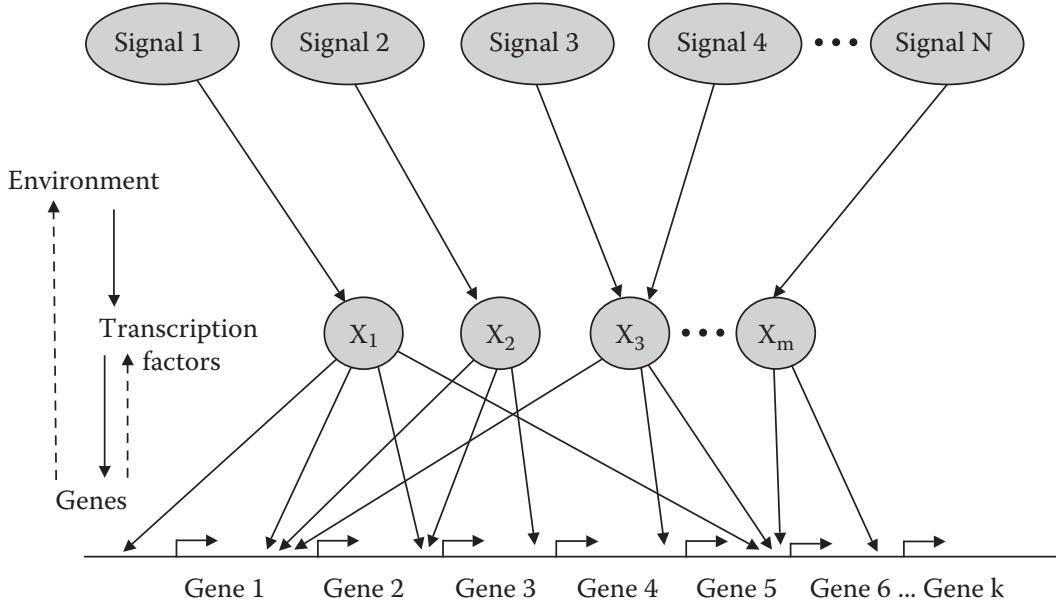


Figure 13.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 13.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

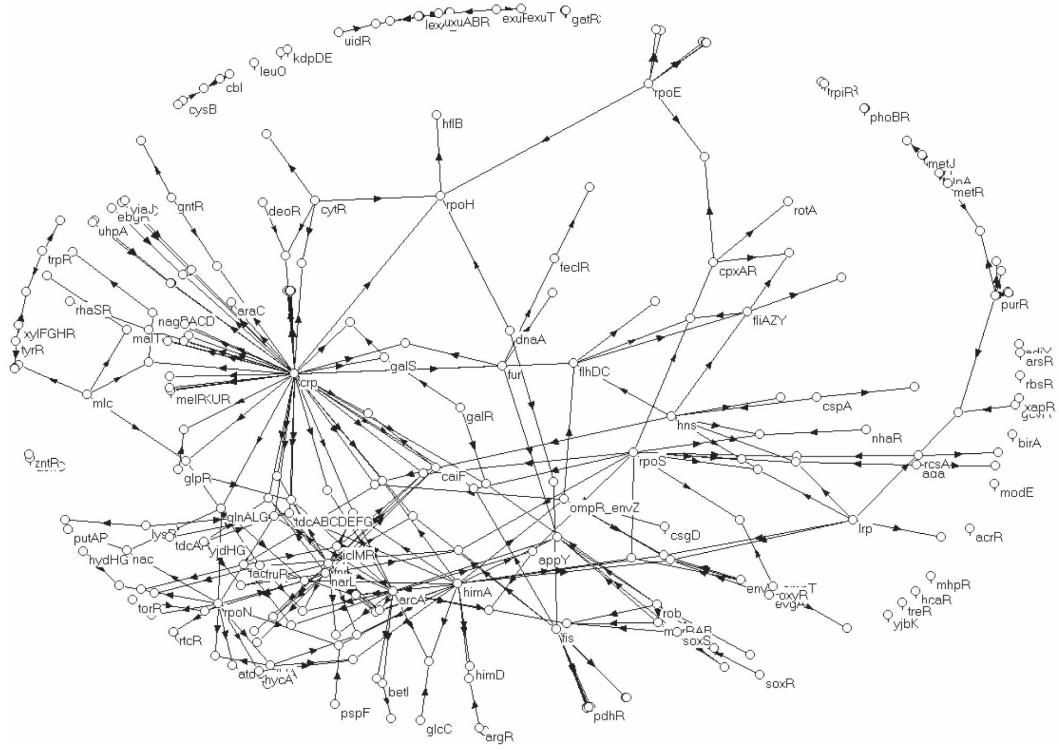


Figure 13.2 Figure caption goes here. Figure caption goes here.

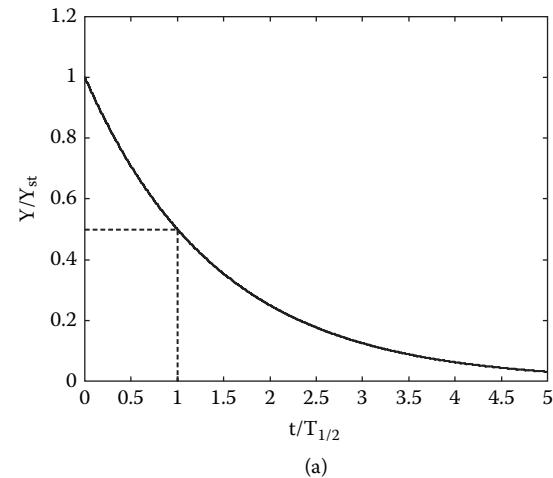
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -1) - 2(3, 5) = (0, 0).$$

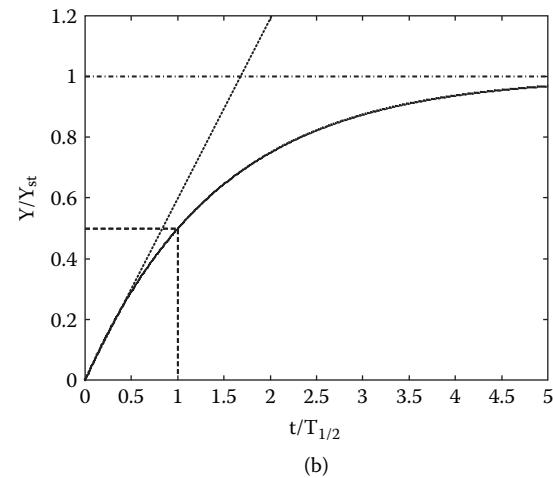
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 13.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 13.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 13.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P13.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 13.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P13.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P13.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P13.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P13.5})$$

- 13.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{C}\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

14.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

14.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 14.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 14.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (14.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (14.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 14.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 14.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

14.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

14.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

14.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

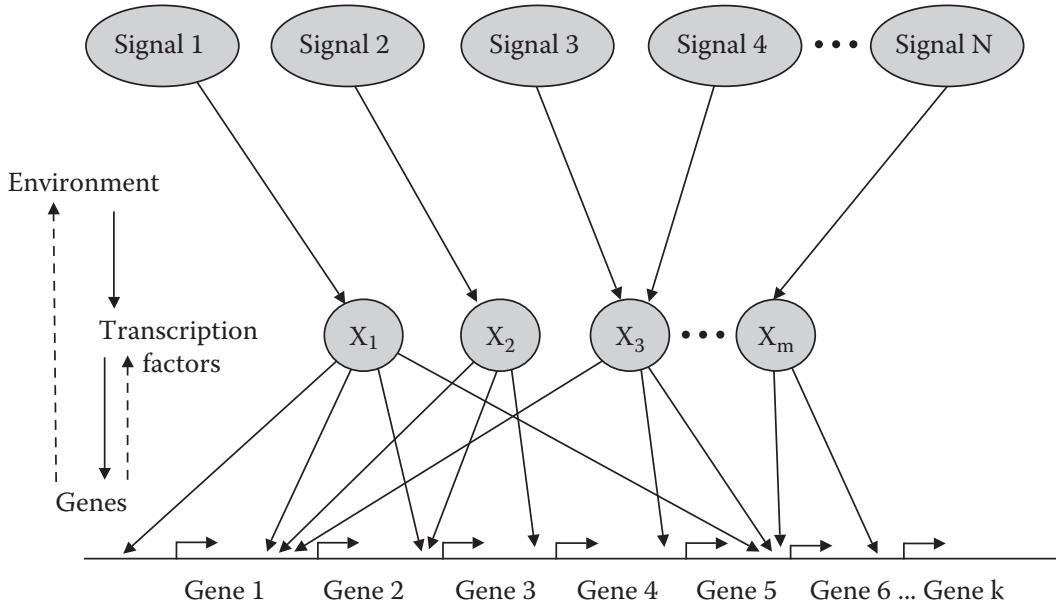


Figure 14.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 14.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

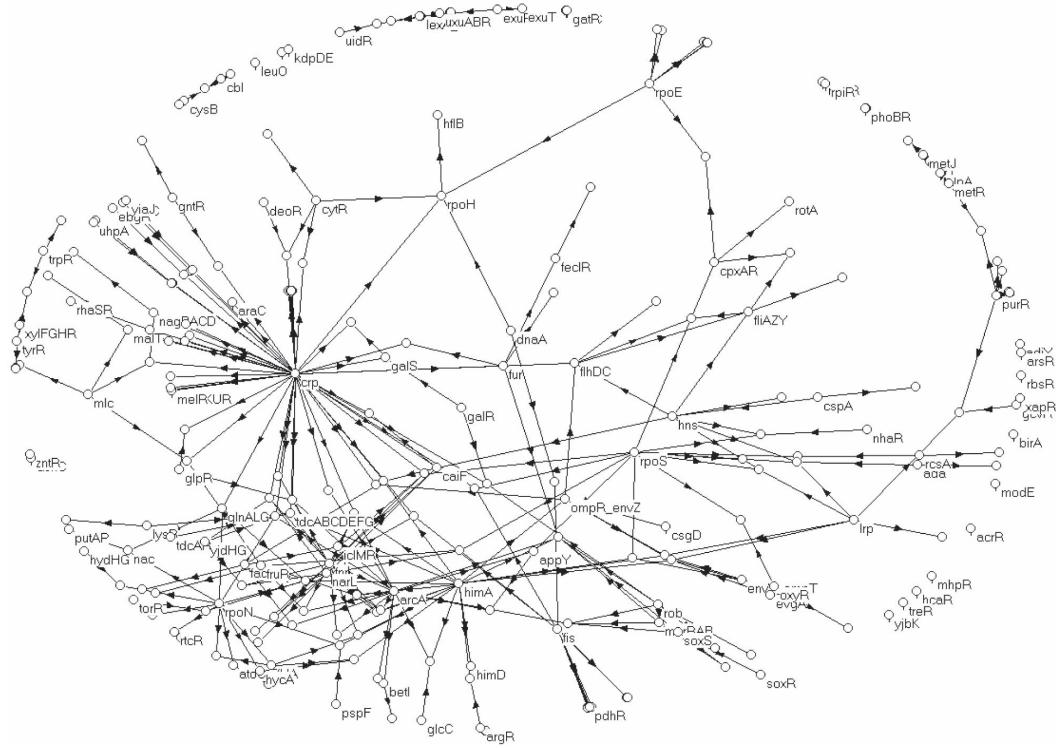


Figure 14.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

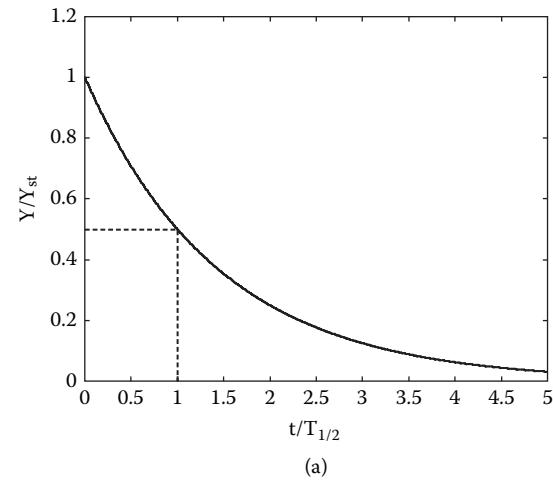
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

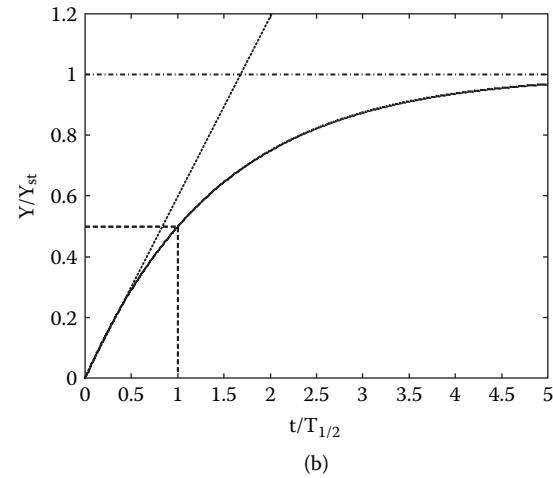
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 14.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 14.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 14.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P14.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 14.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P14.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P14.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P14.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P14.5})$$

- 14.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{C}\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

15.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

15.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 15.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 15.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (15.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (15.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 15.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 15.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

15.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

15.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

15.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

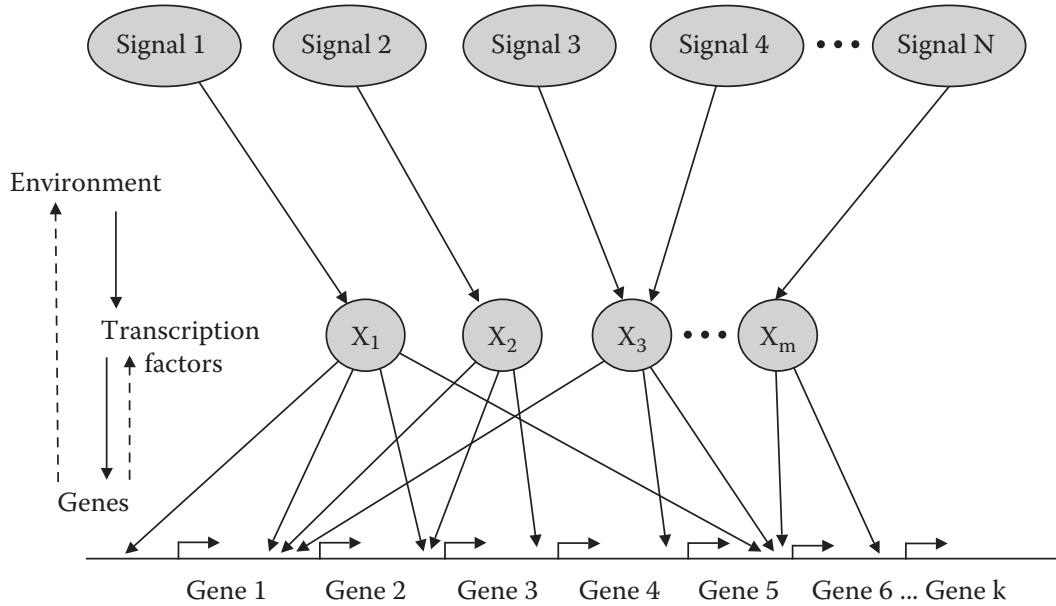


Figure 15.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 15.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

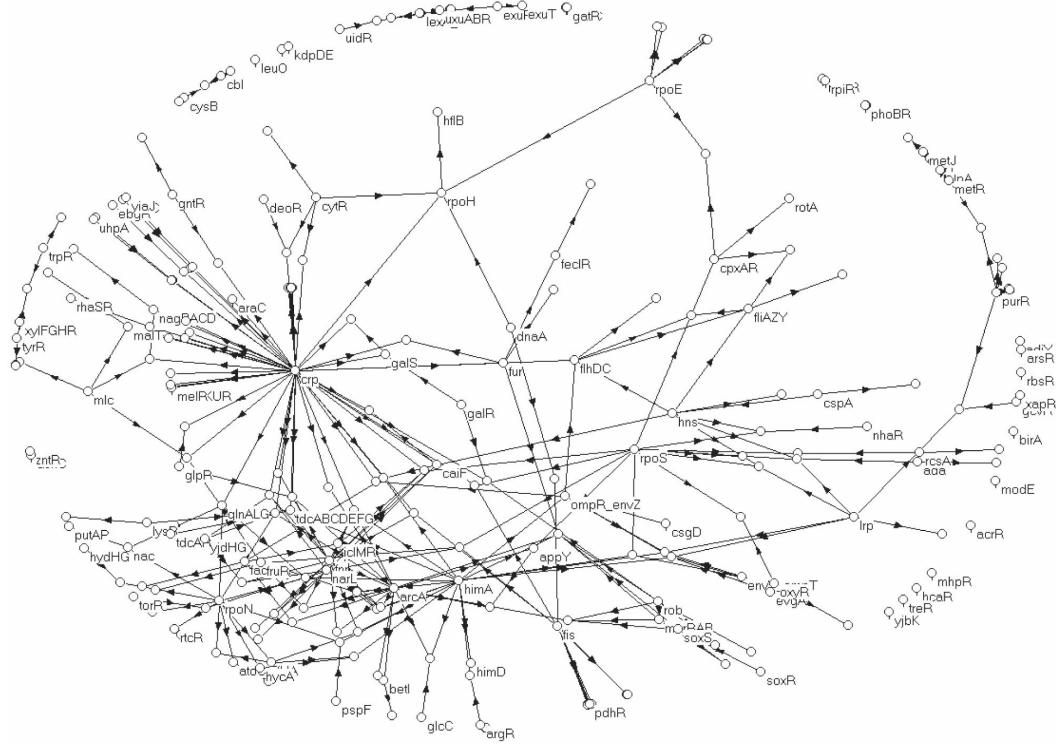


Figure 15.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

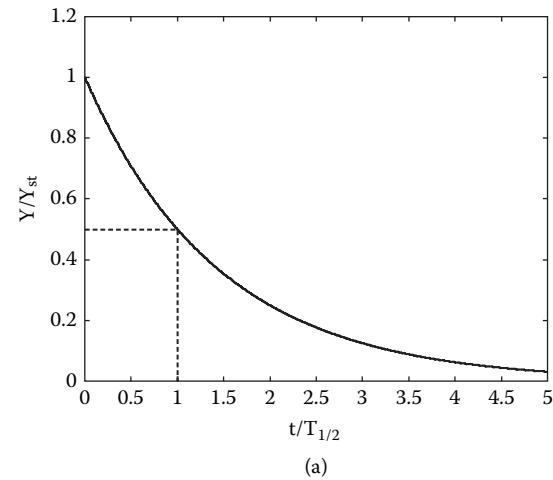
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

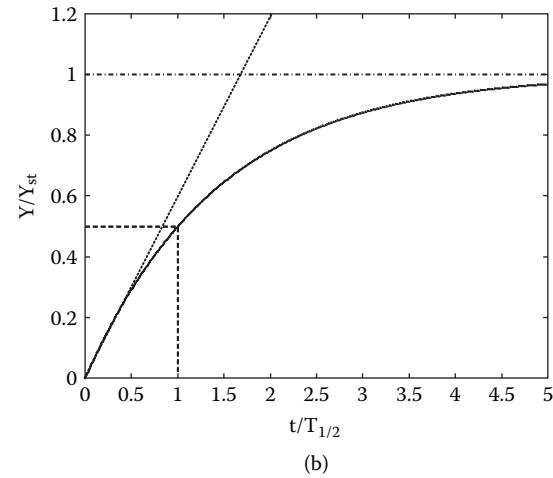
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 15.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 15.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 15.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P15.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 15.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P15.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P15.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P15.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P15.5})$$

- 15.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{C}\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

16.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

16.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 16.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 16.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (16.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (16.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 16.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 16.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

16.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

16.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

16.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

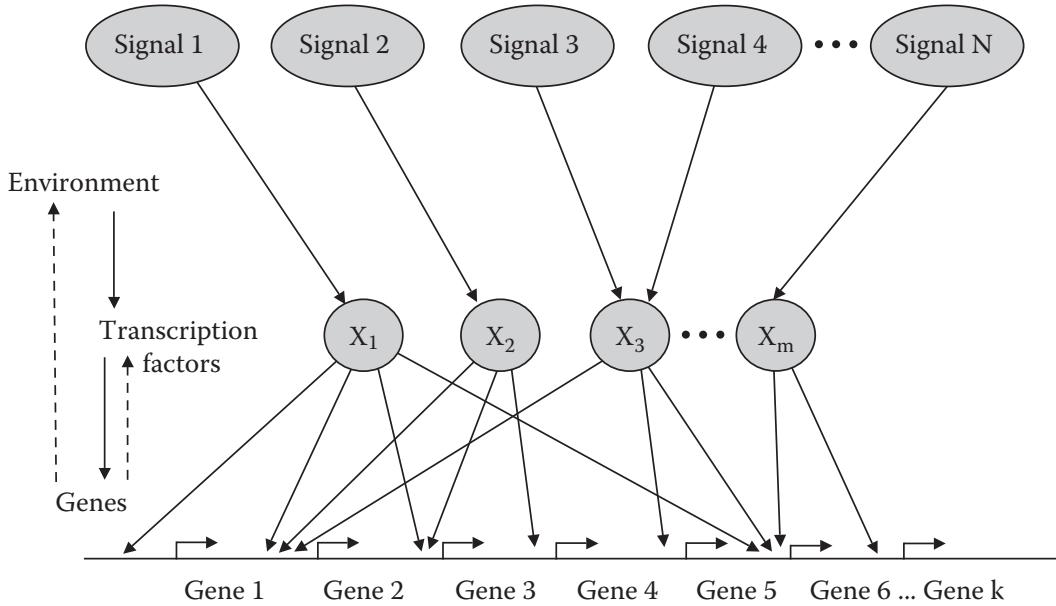


Figure 16.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 16.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

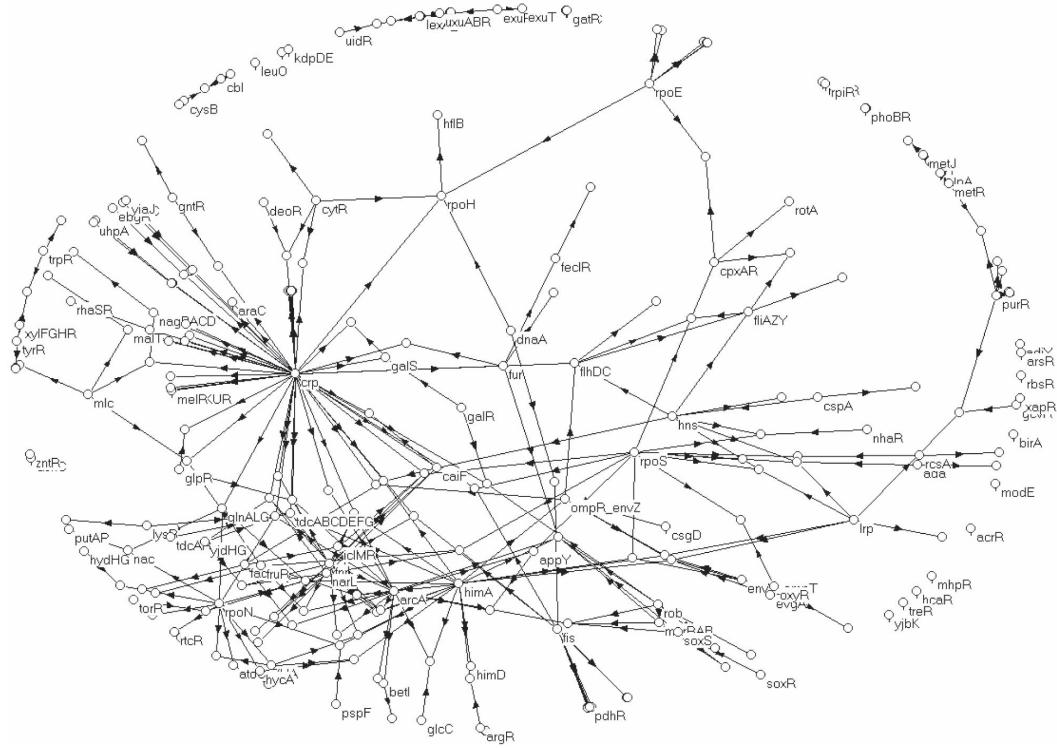


Figure 16.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

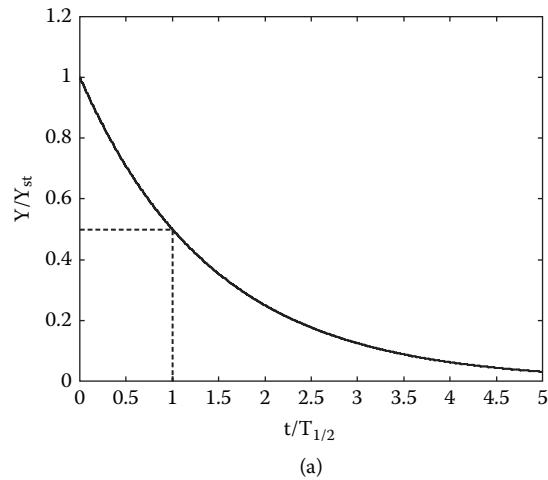
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

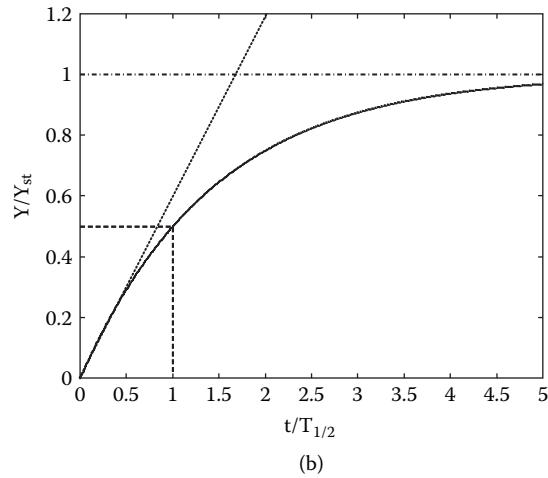
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 16.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 16.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 16.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P16.1})$$

The solution of such an equation is generally $Y = C_1 + r C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 16.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P16.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P16.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P16.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P16.5})$$

- 16.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash\{\mathbb{e}\}$
- $\mathbb{C}\mathbb{E} = \backslash\{\mathbb{c}\mathbb{e}\}$
- $\emptyset = \backslash\circ$

17.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

17.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 17.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

"A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action."

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure *k*-anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table *PT* and a set of tables T_1, \dots, T_n , where *PT* contains a person identifier *Pid* and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in *QID*, and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same *QID* with o . It is important to emphasize that the *k*-anonymization is applied at the *record owner* level, not at the *record* level in traditional *k*-anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 17.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (17.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (17.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 17.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 17.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

17.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

17.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

17.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

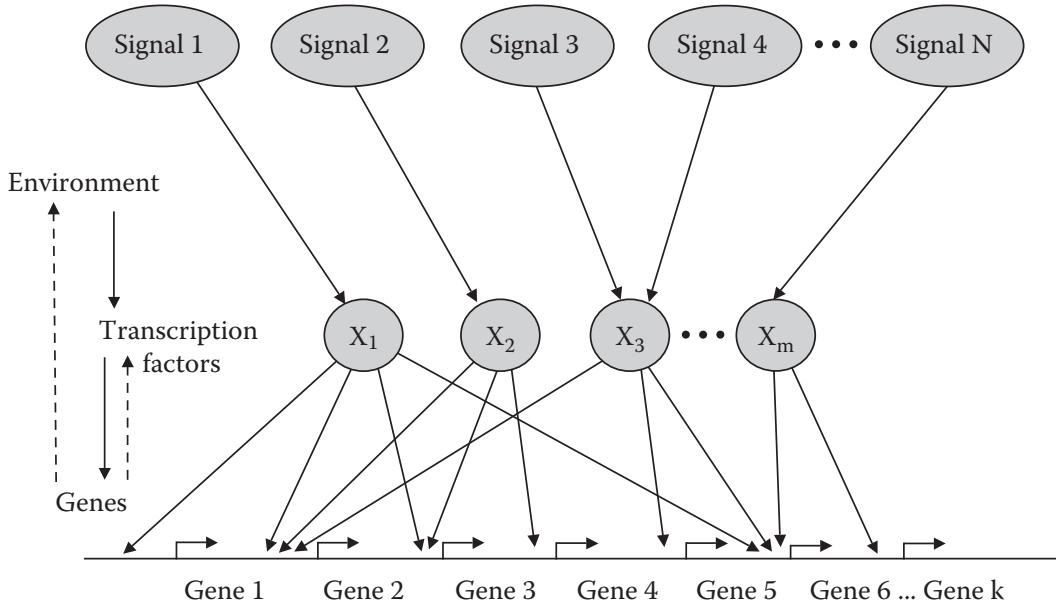


Figure 17.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 17.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

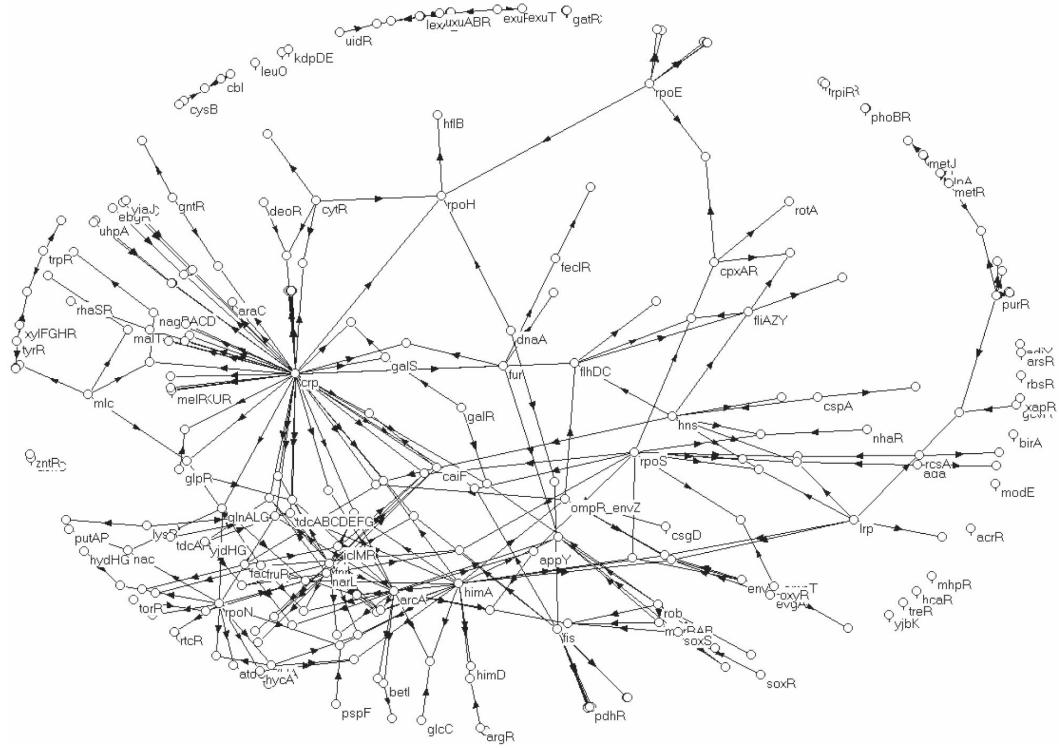


Figure 17.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

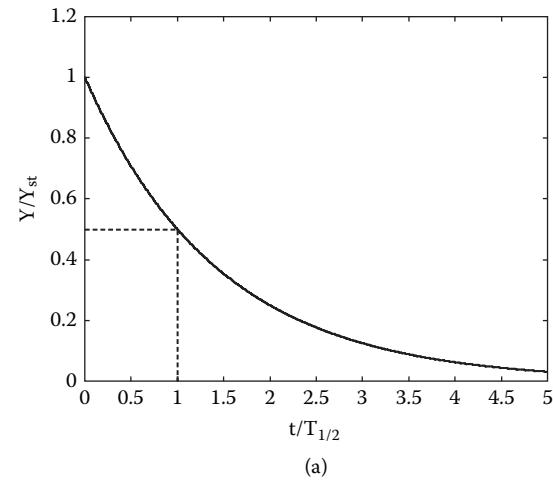
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

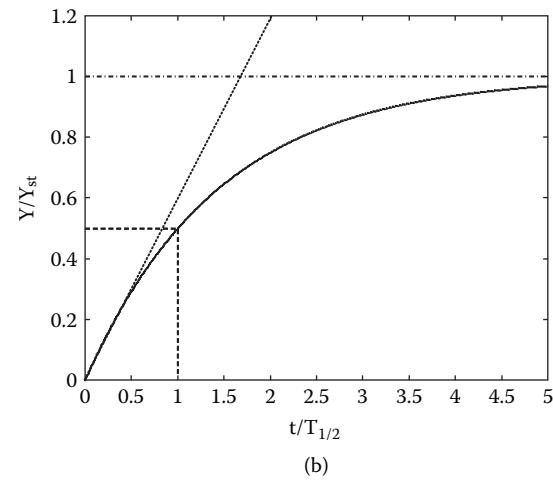
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 17.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 17.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 17.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P17.1})$$

The solution of such an equation is generally $Y = C_1 + r C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 17.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P17.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P17.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P17.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P17.5})$$

- 17.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{C}\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

18.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

18.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 18.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 18.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (18.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (18.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 18.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 18.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

18.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

18.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

18.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

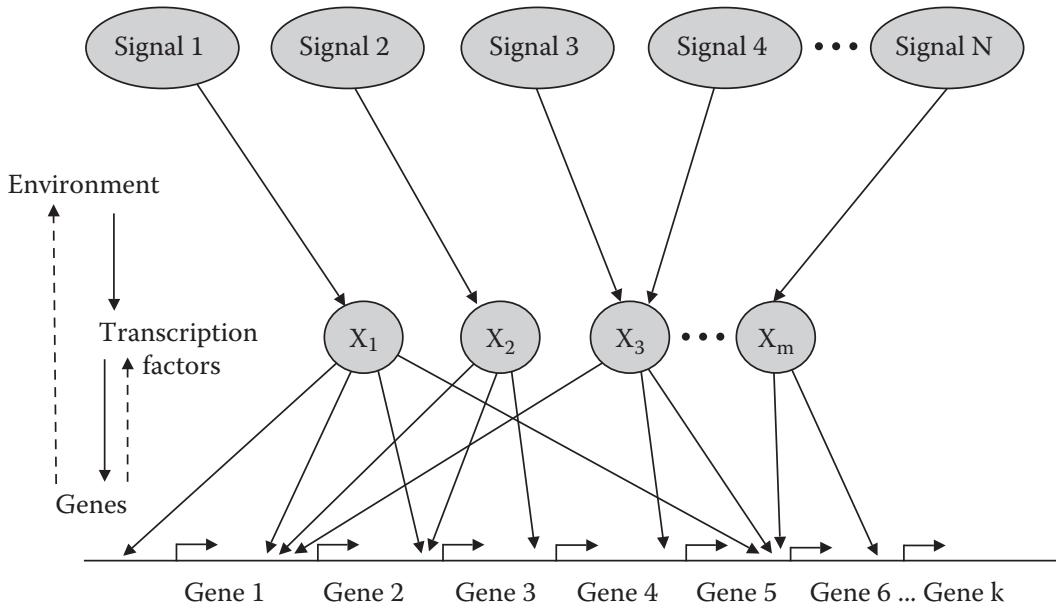


Figure 18.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 18.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

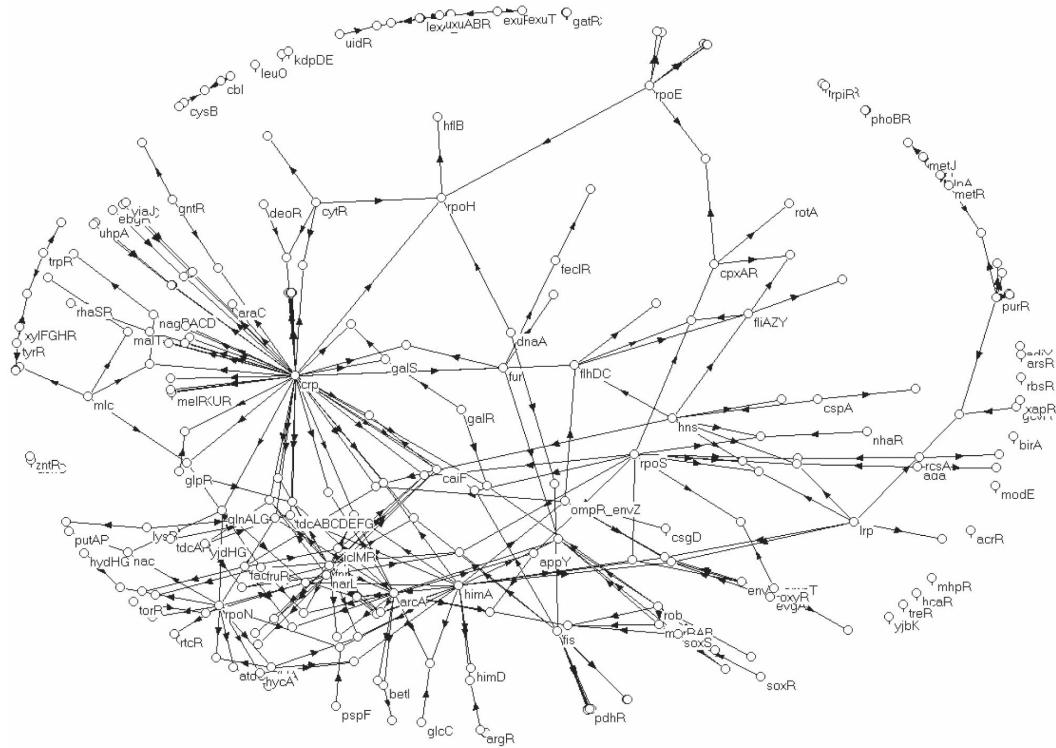


Figure 18.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

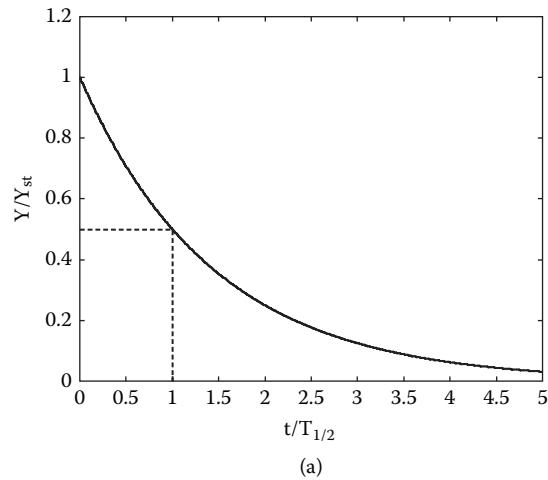
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

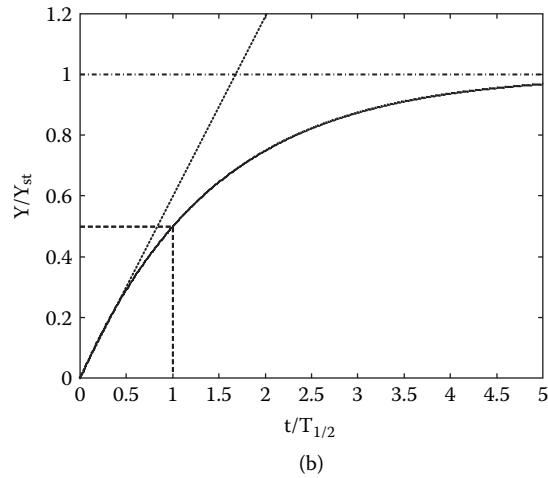
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 18.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 18.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 18.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P18.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 18.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P18.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P18.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P18.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P18.5})$$

- 18.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{C}\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

19.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

19.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 19.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 19.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (19.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (19.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 19.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 19.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

19.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

19.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

19.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

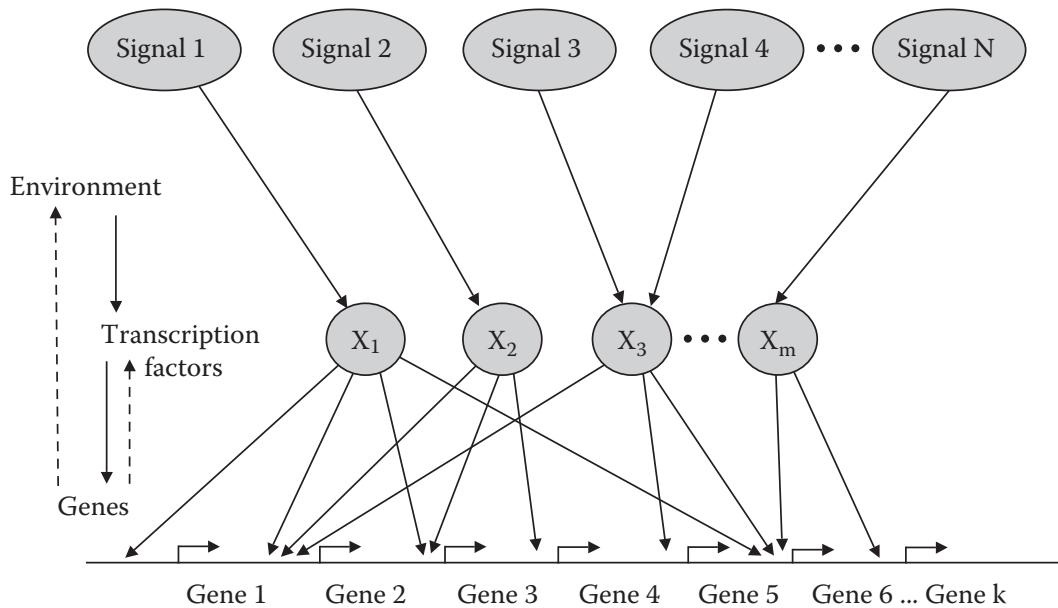


Figure 19.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 19.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

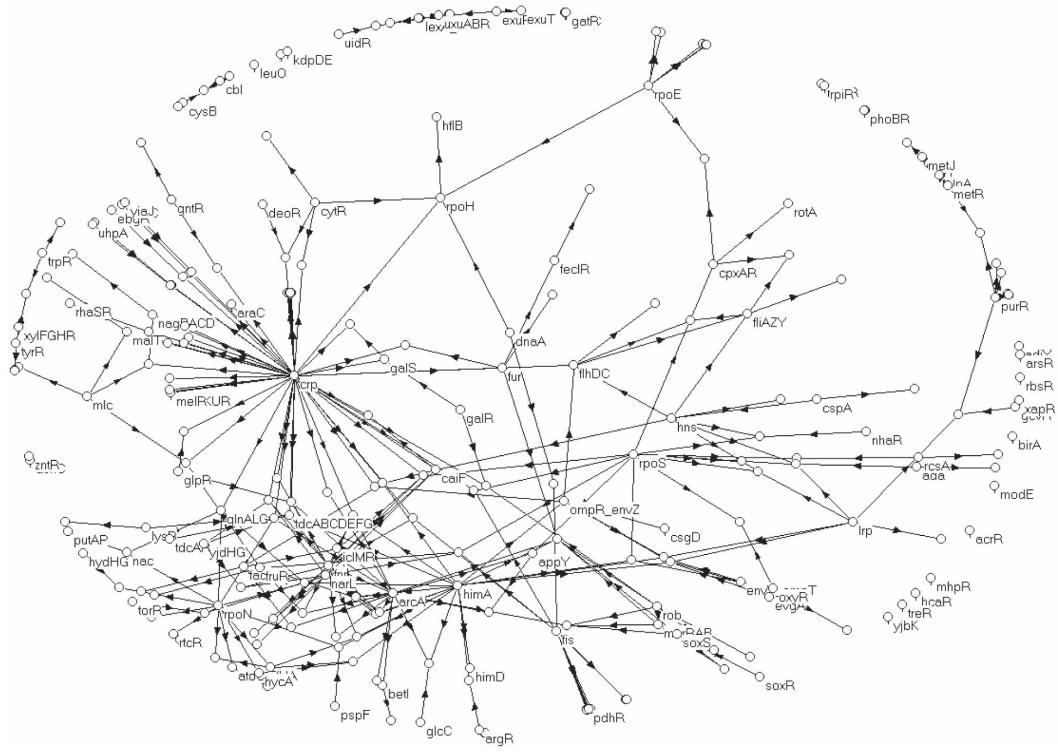


Figure 19.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

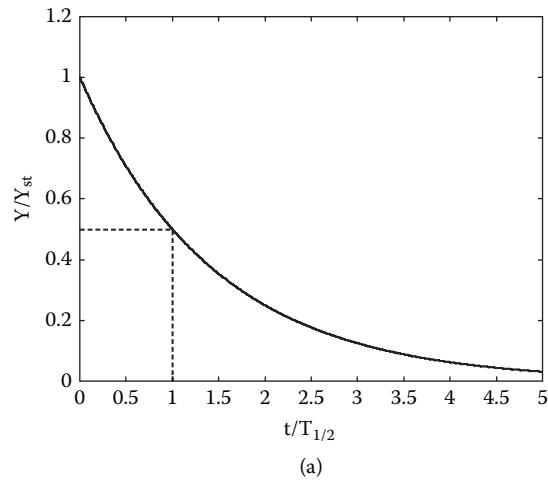
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

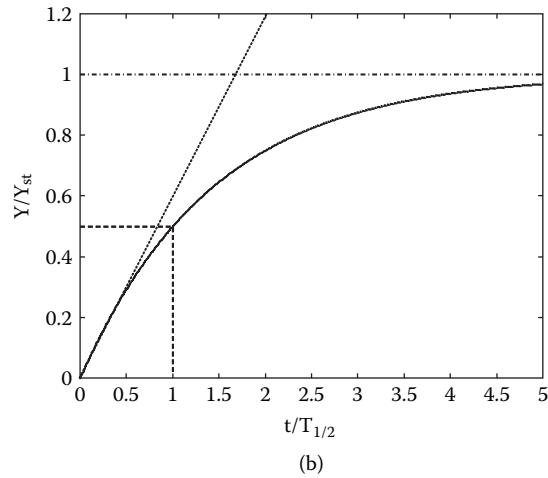
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set* of V provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors* of V provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 19.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 19.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 19.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P19.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 19.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P19.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P19.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P19.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P19.5})$$

- 19.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

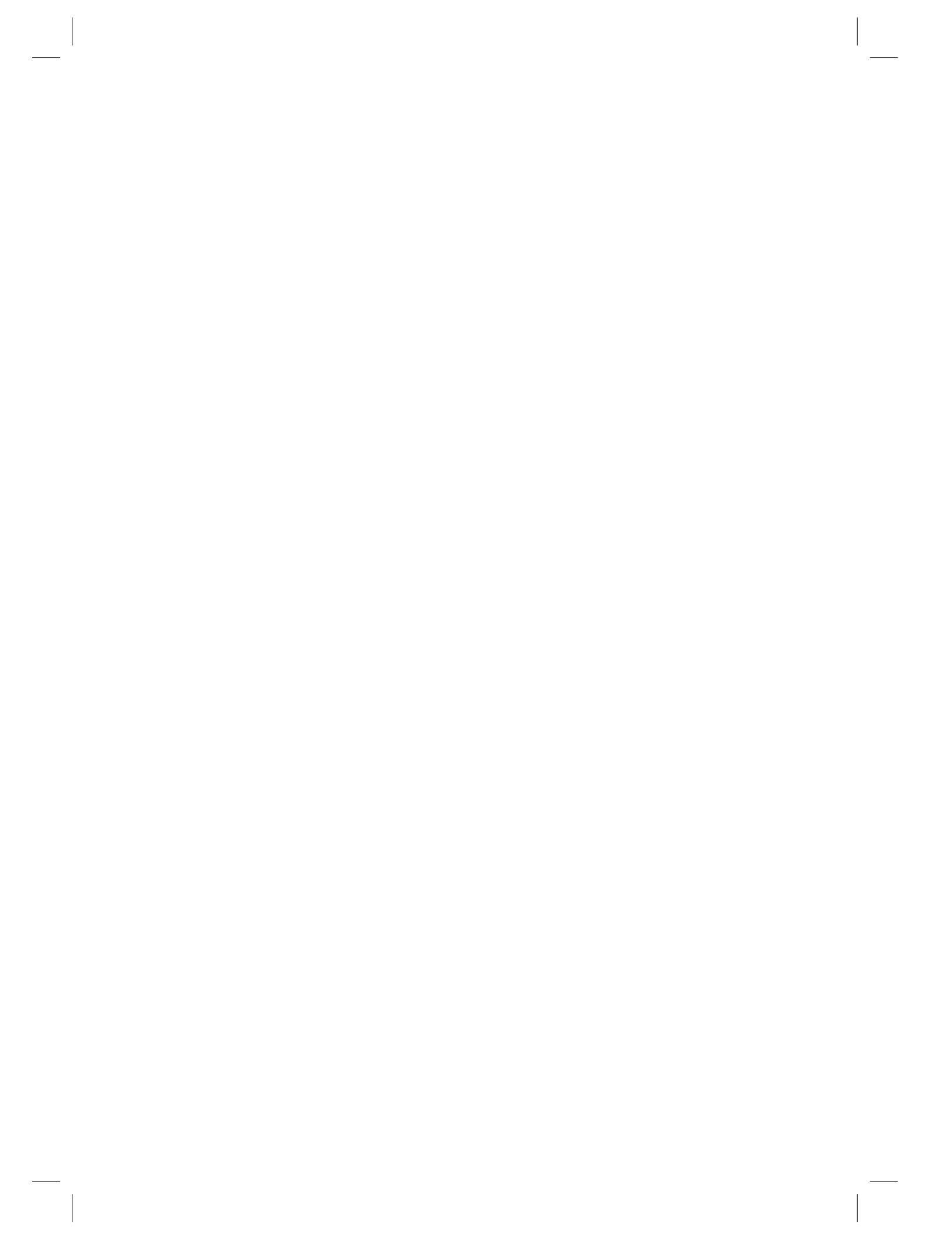
FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

20.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

20.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 20.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 20.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (20.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (20.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 20.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 20.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

20.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

20.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

20.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

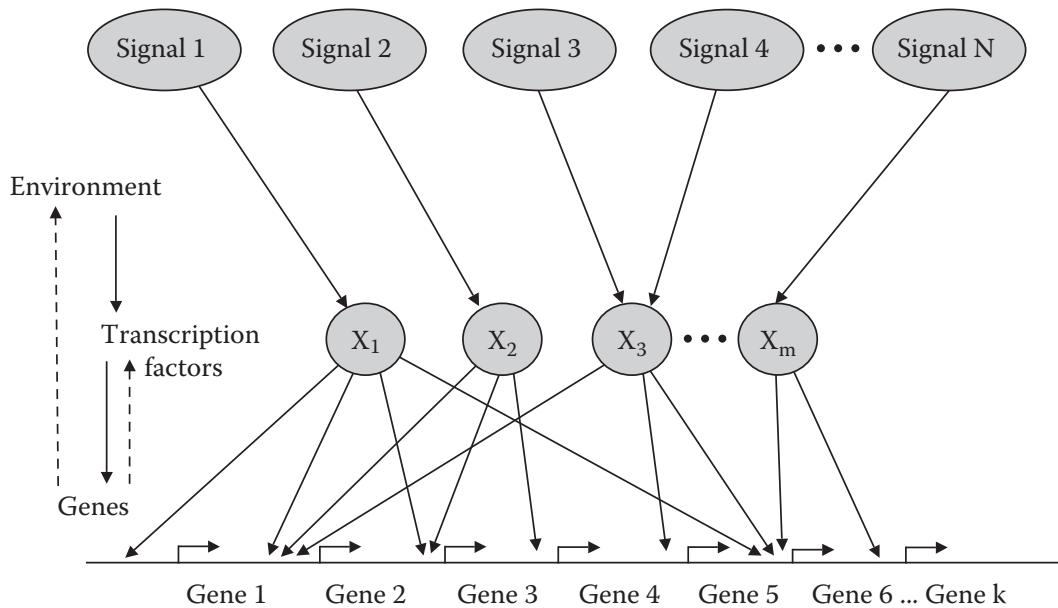


Figure 20.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 20.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

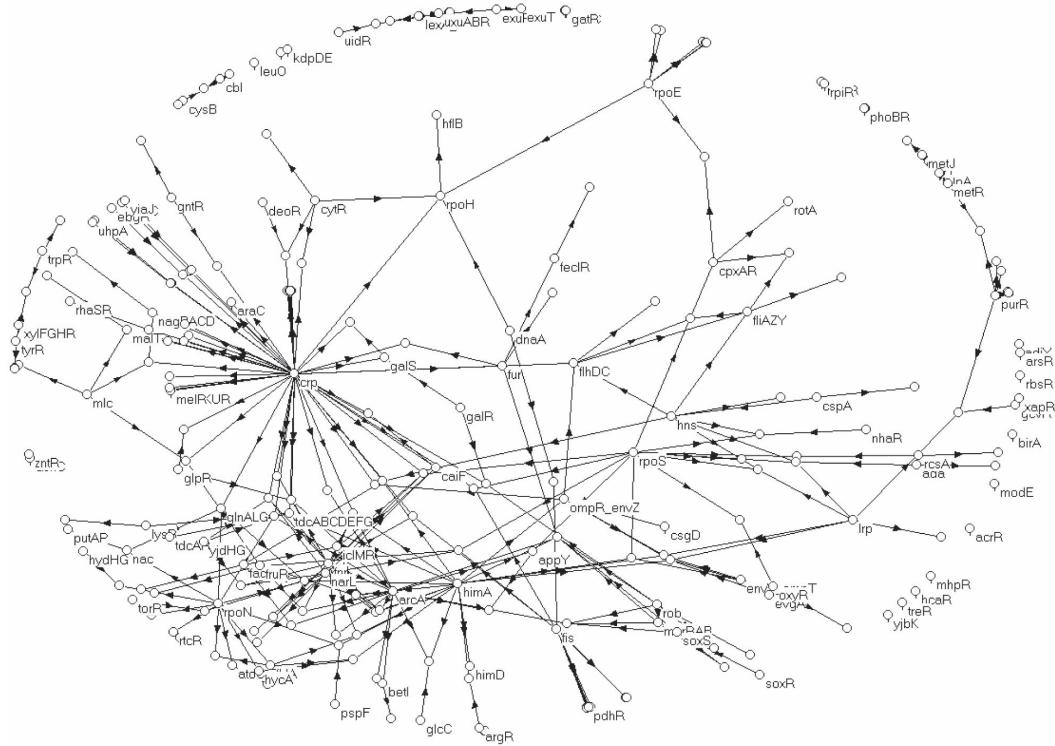


Figure 20.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here. Figure caption goes here.

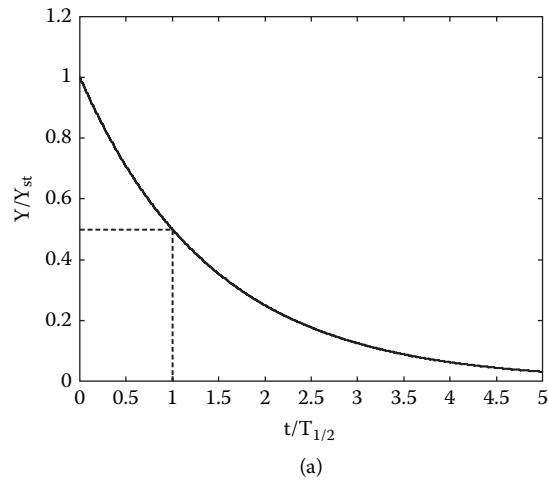
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

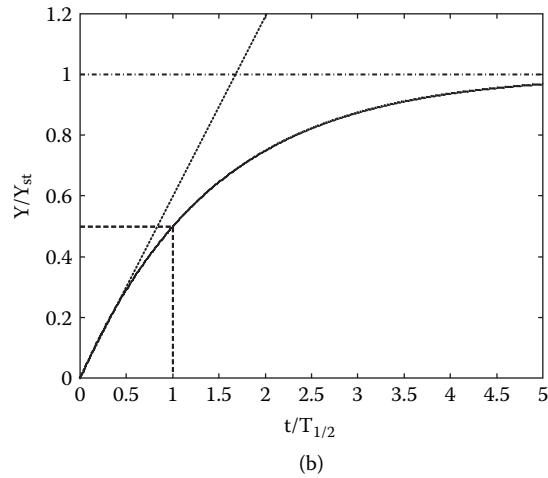
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 20.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 20.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 20.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P20.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 20.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P20.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P20.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P20.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P20.5})$$

- 20.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Introduction to Image Processing and the MATLAB Environment

A COMPONENT PART for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

- Checking various symbols in the current package configuration
- _ — _
- _ — \textunderscore
- \ — \textbackslash
- TM — \texttrademark
- [®] — \textregistered
- < — \textless
- > — \textgreater
- etc.
- Currently OK symbols with this package configuration
- å — \aa
- ø — \r{a}
- ÿ — \u{u}

- $\mathbb{E} = \backslash' \{e\}$
- $\mathbb{E} = \backslash' \{e\}$
- $\emptyset = \backslash o$

21.1 INTRODUCTION

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

21.1.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line.

TABLE 21.1 Comparison of MATLAB and C Code for Simple Matrix Operations

| Operations | Part of C Code | Hor. fts. | Ver. fts. |
|---------------------|----------------|-----------|-----------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi ^a | 46, 281 | 6, 898 | 5, 400 |
| Keybrd ^b | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty [4]? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty? A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

“A Process is a structured, measured set of activities designed to produce a specific output for a particular customer or market—A process is thus a specific ordering of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs: a structure for action.”

Thomas Davenport
Senior Adjutant to the Junior Marketing VP

MultiRelational k -Anonymity. Most works on k -anonymity focus on anonymizing a single data table; however, a real-life [2] database usually contains mul-

multiple relational tables. This has proposed a privacy model called *MultiR k-anonymity* to ensure k -anonymity on multiple relational tables. Their model assumes that a relational database contains a person-specific table PT and a set of tables T_1, \dots, T_n , where PT contains a person identifier Pid and some sensitive attributes, and T_i , for $1 \leq i \leq n$, contains some foreign keys, some attributes in QID , and sensitive attributes. The general privacy notion is to ensure that for each record owner o contained in the join of all tables $PT \bowtie T_1 \bowtie \dots \bowtie T_n$, there exists at least $k - 1$ other record owners share the same QID with o . It is important to emphasize that the k -anonymization is applied at the *record owner* level, not at the *record* level in traditional k -anonymity. This idea is similar to (X, Y) -anonymity, where $X = QID$ and $Y = \{Pid\}$.

1. Factual knowledge (“knowing that”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - (a) Conceptual knowledge (“knowing why”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - i. Procedural knowledge (“knowing what”) One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage, attribute linkage*.
 - ii. Spatial knowledge (“knowing what”)

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker’s background knowledge. Below, the term “victim” refers to the record owner being linked. We can broadly classify linking models to two families.

A component part for an electronic item is [3] manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C

TABLE 21.2 Now we are engaged (a_g^g) (a_g^a) in a great civil war, testing whether that nation, or any nation so conceived.

| | | | |
|--------|---------|--------|--------|
| Ball | 19, 221 | 4, 598 | 3, 200 |
| Pepsi | 46, 281 | 6, 898 | 5, 400 |
| Keybrd | 27, 290 | 2, 968 | 3, 405 |
| Pepsi | 14, 796 | 9, 188 | 3, 209 |

20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively.

In most literature on PPDP, they [5] consider a more relaxed, yet more practical, notion of privacy protection by assuming limited attacker's background knowledge. Below, the term "victim" refers to the record owner being linked. We can broadly classify linking models to two families.

- Factual knowledge ("knowing that") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - Conceptual knowledge ("knowing why") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Procedural knowledge ("knowing what") One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*.
 - * Spatial knowledge ("knowing what")

One family considers a privacy threat occurs when an attacker is able to link a record owner to a record in a published data table, to a sensitive attribute in a published data table, or to the published data table itself. We call them *record linkage*, *attribute linkage*, and *table linkage*, respectively. In all types of linkages, we assume that the attacker knows the *QID* of the victim. In record and attribute linkages, we further assume that the attacker knows the presence of the victim's record in the released table, and seeks to identify the victim's record and/or sensitive information from the table [9]. In table linkage, the attack seeks to determine the present or absent of the victim's record in the released table. A data table is considered to privacy preserved if the table can effectively prevent the attacker from successfully performing these types of linkages on the table [6]. Sections 21.1-21.2 study this family of privacy models.

$$\text{var} \hat{\Delta} = \sum_{j=1}^t \sum_{k=j+1}^t \text{var} (\hat{\alpha}_j - \hat{\alpha}_k) = \sum_{j=1}^t \sum_{k=j+1}^t \sigma^2 (1/n_j + 1/n_k). \quad (21.1)$$

An obvious measure of imbalance is just the difference in the number of times the two treatments are allocated

$$D_n = \mathcal{M}|n_A - n_B|. \quad (21.2)$$

For rules such as deterministic allocation, for which the expected value of this difference can be calculated, we obtain the population value \mathcal{D}_n .

Theorem 21.1 Let m be a prime number. With the addition and multiplication as defined above, Z_m is a field.

Proof 21.1 Most of the proof of this theorem is routine. It is clear that $0 \in Z_m$ and $1 \in Z_m$ are the zero element and identity element. If $a \in Z_m$ and $a \neq 0$, then $m - a$ is the additive inverse of a . If $a \in Z_m$ and $a \neq 0$, then the greatest common divisor of a and m is 1, and hence there exist integers s and t such that $sa + tm = 1$. Thus $sa = 1 - tm$ is congruent to 1 modulo m . Let s^* be the integer in Z_m congruent to s modulo m . Then we also have $s^*a \equiv 1 \pmod{m}$. Hence s^* is the multiplicative inverse of a modulo m . Verification of the rest of the field properties is now routine.

21.2 RECORD LINKAGE MODEL

In the privacy attack of *record linkage*, some value qid on QID identifies a small number of records in the released table T , called a *group*. If the victim's QID matches the value qid , the victim is vulnerable to being linked to the small number of records in the group [7]. In this case, the attacker faces only a small number of possibilities for the victim's record, and with the help of additional knowledge, there is a chance that the attacker could uniquely identify the victim's record from the group.

21.2.1 A Component Part

A component part for an electronic item is manufactured at one of three different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

21.2.1.1 H3 A Component Part

A component part for an electronic item is manufactured at one of three [8] different factories, and then delivered to the main assembly line. Of the total number supplied, factory A supplies 50%, factory B 30%, and factory C 20%. Of the components manufactured at factory A, 1% are faulty and the corresponding proportions for factories B and C are 4% and 2% respectively. A component is picked at random from the assembly line. What is the probability that it is faulty?

H4 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

H5 A Component Part A fundamental notion [9] is that of a subspace of F^n . Let V be a nonempty subset of F^n . Then V is a *subspace* of F^n provided V is closed under vector addition and scalar multiplication, that is,

1. For all u and v in V , $u + v$ is also in V .

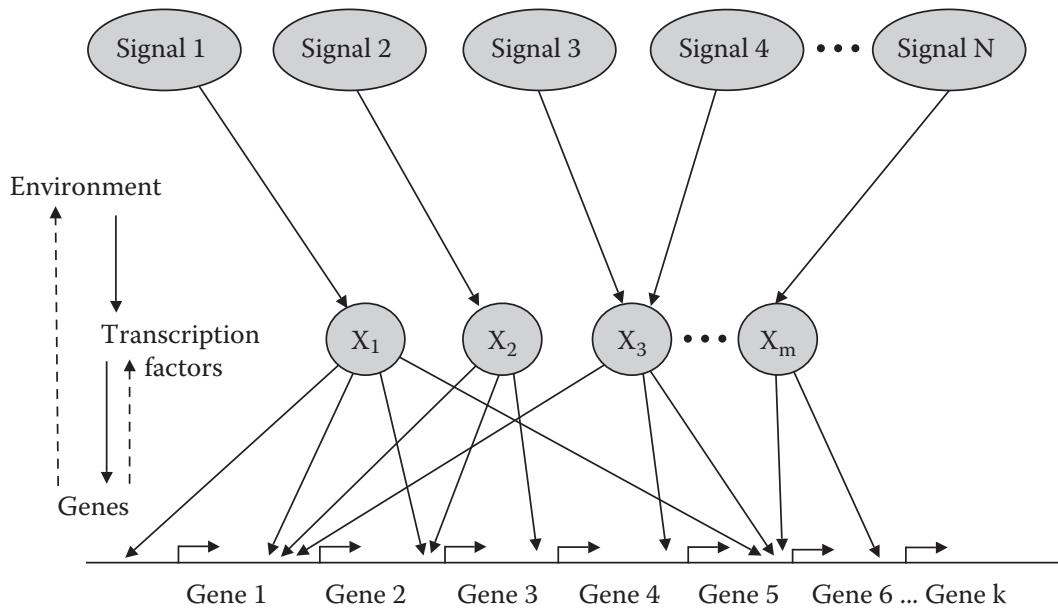


Figure 21.1 Figure caption goes here.

2. For all u in V and c in F , cu is in V .

Let u be in the subspace V . Because $0u = 0$, it follows that the zero vector is in V . Similarly, $-u$ is in V for all u in V . A simple example of a subspace of F^n is the set of all vectors $(0, a_2, \dots, a_n)$ with first coordinate equal to 0. The zero vector itself is a subspace.

Definition 21.1 Let $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ be vectors in F^n , and let c_1, c_2, \dots, c_m be scalars. Then the vector

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)}$$

is called a *linear combination* of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. If V is a subspace of F^n , then V is closed under vector addition and scalar multiplication, and it follows easily by induction that a linear combination of vectors in V is also a vector in V . Thus *subspaces are closed under linear combinations*; in fact, this can be taken as the defining property of subspaces. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ *span* V (equivalently, form a *spanning set* of V) provided every vector in V is a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. The zero vector can be written as a linear combination of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ with all scalars equal to 0; this is a *trivial linear combination*. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are *linearly dependent* provided there are scalars c_1, c_2, \dots, c_m , not all of which are zero, such that

$$c_1u^{(1)} + c_2u^{(2)} + \dots + c_mu^{(m)} = 0,$$

that is, the zero vector can be written as a *nontrivial linear combination*

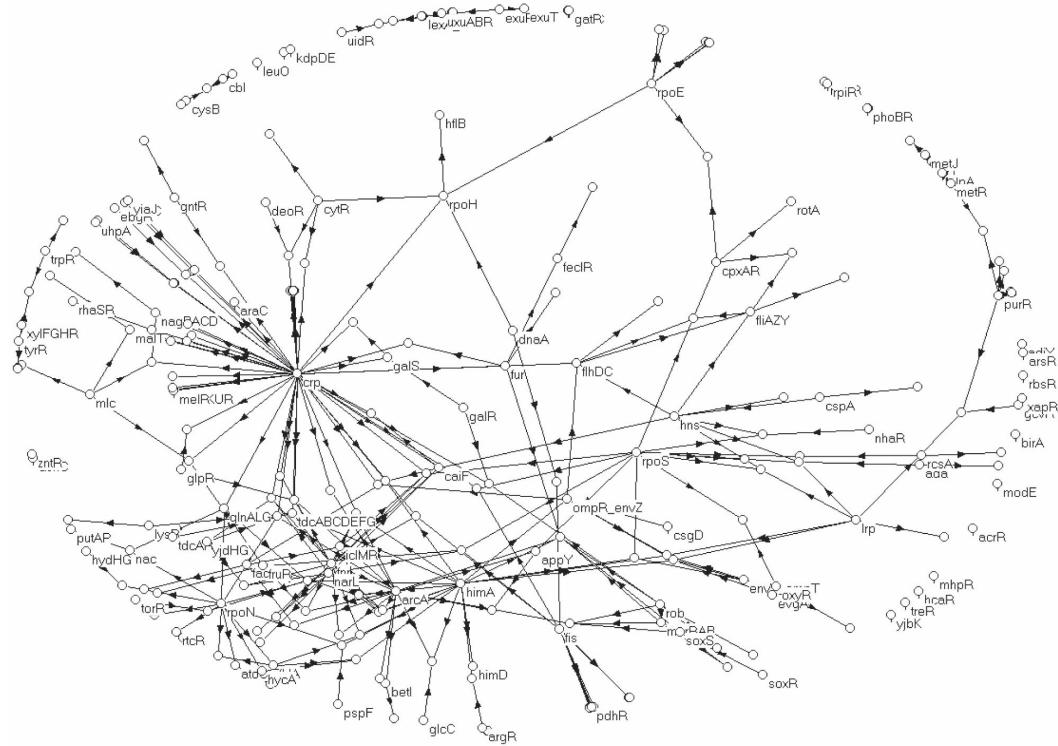


Figure 21.2 Figure caption goes hereq. Figure caption goes here. Figure caption goes here.

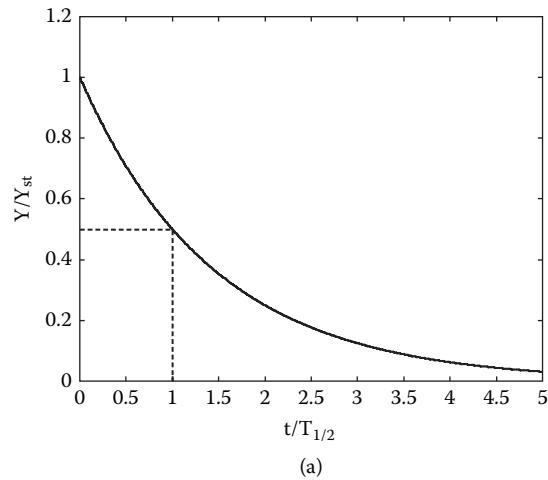
of $u^{(1)}, u^{(2)}, \dots, u^{(m)}$. For example, the vectors $(1, 4), (3, -1)$, and $(3, 5)$ in \mathbb{R}^2 are linearly dependent since

$$3(1, 4) + 1(3, -2) - 2(3, 5) = (0, 0).$$

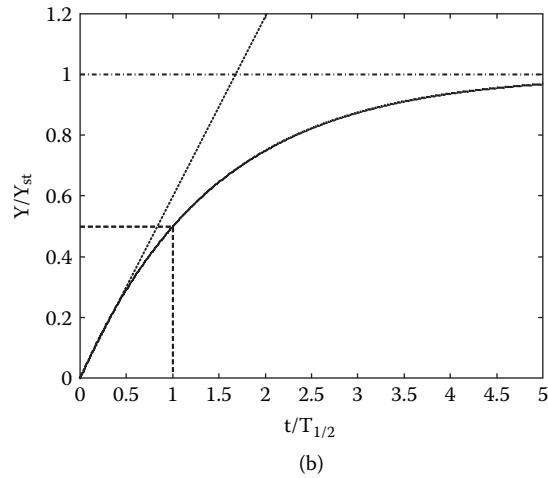
Vectors are *linearly independent* provided they are not linearly dependent. The vectors $u^{(1)}, u^{(2)}, \dots, u^{(m)}$ are a *basis* of V provided they are linearly independent and span V . By an *ordered basis* we mean a basis in which the vectors of the basis are listed in a specified order; to indicate that we have an ordered basis we write $(u^{(1)}, u^{(2)}, \dots, u^{(m)})$. A spanning set S of V is a *minimal spanning set of V* provided that each set of vectors obtained from S by removing a vector is not a spanning set for V . A linearly independent set S of vectors of V is a *maximal linearly independent set of vectors of V* provided that for each vector w of V that is not in S , $S \cup \{w\}$ is linearly dependent (when this happens, w must be a linear combination of the vectors in S). \square

In addition to matrix addition, subtraction, and multiplication, there is one additional operation that we define now. It's perhaps the simplest of them all. Let $A = [a_{ij}]$ be an m by n matrix and let c be a number [3]. Then the matrix $c \cdot A$, or simply cA , is the m by n matrix obtained by multiplying each entry of A by c :

$$cA = [ca_{ij}].$$



(a)



(b)

Figure 21.3 The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach, the white ones correspond to shortfall risk.

The matrix cA is called a *scalar multiple* of A .

The term reliability usually refers to the probability that a component or system will operate satisfactorily either at any particular instant at which it is required or for a certain length of time. Fundamental to quantifying reliability is a knowledge of how to define, assess and combine probabilities [1]. This may hinge on identifying the form of the variability which is inherent in most processes. If all components had a fixed known lifetime there would be no need to model reliability.

EXERCISES

-
- 21.1 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.2 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.3 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.4 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.5 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.6 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.7 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- 21.8 *A change in production rate.* A gene Y with simple regulation is produced at a constant rate β_1 . The production rate suddenly shifts to a different rate β_2 .
- Calculate and plot the gene product concentration $Y(t)$.
 - What is the response time (time to reach halfway between the steady states)?

Solution (for part a):

- Let us mark the time when the shift occurs as $t = 0$. Before the shift, Y reaches steady state at a level $Y(t = 0) = Y_{st} = \beta_1/\alpha$. After the shift,

$$dY/dt = \beta_2 - \alpha Y \quad (\text{P21.1})$$

The solution of such an equation is generally $Y = C_1 + C_2 e^{-\alpha t}$, where the constants C_1 and C_2 need to be determined so that $Y(t = 0) = \beta_1/\alpha$, and Y at long times reaches its new steady state, β_2/α . This yields the following sum of an exponential and a constant:

$$Y(t) = \beta_2 / \alpha + (\beta_1 / \alpha - \beta_2 / \alpha)e^{-\alpha t} \quad (\text{P2.2})$$

Take the derivative with respect to time, dY/dt , and verify that Equation P2.1 is fulfilled.

- 21.9 *mRNA dynamics.* In the main text, we considered the activation of transcription of a gene (mRNA production) and used a dynamical equation to describe the changes in the concentration of the gene product, the protein Y. In this equation, $dY/dt = \beta - \alpha Y$, the parameter β describes the rate of protein production. In reality, mRNA needs to be translated to form the protein, and mRNA itself is also degraded by specific enzymes.
- Derive dynamical equations for the rate of change of mRNA and the rate of change of the protein product, assuming that mRNA is produced at rate β_m and degraded at rate α_m , and that each mRNA produces on average p protein molecules per unit time. The protein is degraded/diluted at rate α .
 - Note that mRNA is often degraded at a much faster rate than the protein product $\alpha_m \gg \alpha$. Can this be used to form a quasi-steady-state assumption that mRNA levels are at steady state with respect to slower processes? What is the effective protein production rate β in terms of β_m , α_m , and p? What would be the response time if the mRNA lifetime were much longer than the protein lifetime?

Solution:

- a. The dynamic equation for the concentration of mRNA of gene Y, Y_m , is:

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P21.2})$$

The dynamical equation for the protein product is due to production of p copies per mRNA and degradation/dilution at rate α :

$$dY_m / dt = \beta_m - \alpha_m Y_m \quad (\text{P21.3})$$

$$dY / dt = p Y_m - \alpha Y \quad (\text{P21.4})$$

- b. In the typical case that mRNA degradation is faster than the degradation/dilution of the protein product, we can assume that Y_m reaches steady state quickly in comparison to the protein levels. The reason is that the typical time for the mRNA to reach steady state is the response time $\log(2)/\alpha_m$, which is much shorter than the protein response time $\log(2)/\alpha$ because $\alpha_m \gg \alpha$. The steady-state mRNA level is found by setting $dY_m/dt = 0$ in Equation P2.3, yielding

Using this for Y_m in Equation P2.4 yields the following equation for the protein production rate:

In other words, the effective protein production rate, which is the first term on the right-hand side of the equation, is equal to the steady-state mRNA

level times the number of proteins translated from each mRNA per unit time:

$$\beta = p \beta_m / \alpha_m \quad (\text{P21.5})$$

- 21.10 *Time-dependent production and decay.* A gene Y with simple regulation has a time-dependent production rate $\beta(t)$ and a time-dependent degradation rate $\alpha(t)$. Solve for its concentration as a function of time.
-

Think About It...

Commonly thought of as the first modern computer, ENIAC was built in 1944. It took up more space than an 18-wheeler's tractor trailer and weighed more than 17 Chevrolet Camaros. It consumed 140,000 watts of electricity while executing up to 5,000 basic arithmetic operations per second. One of today's popular microprocessors, the 486, is built on a tiny piece of silicon about the size of a dime.

With the continual expansion of capabilities, computing power will eventually exceed the capacity for human comprehension or human control.

The Information Revolution
Business Week

GLOSSARY

360 Degree Review: Performance review that includes feedback from superiors, peers, subordinates, and clients.

Abnormal Variation: Changes in process performance that cannot be accounted for by typical day-to-day variation. Also referred to as non-random variation.

Acceptable Quality Level (AQL): The minimum number of parts that must comply with quality standards, usually stated as a percentage.

Activity: The tasks performed to change inputs into outputs.

Adaptable: An adaptable process is designed to maintain effectiveness and efficiency as requirements change. The process is deemed adaptable when there is agreement among suppliers, owners, and customers that the process will meet requirements throughout the strategic period.

FURTHER READING

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative auto-regulation speeds the response time of transcription networks. *J. Mol. Biol.*, 323: 785–793.

Savageau, M.A. (1976). *Biochemical Systems Analysis: A study of Function and Design in Molecular Biology*. Addison-Wesley. Chap. 16.

Savageau, M.A. (1974). Comparison of classical and auto-genous systems of regulation in inducible operons. *Nature*, 252: 546–549.



Bibliography

- [1] G. Bontempi and Y. Le Borgne. An adaptive modular approach to the mining of sensor network data. In *Proceedings of the Workshop on Data Mining in Sensor Networks, SIAM SDM*, pages 3–9. SIAM Press, 2005.
- [2] KI Diamantaras and SY Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc. New York, NY, USA, 1996.
- [3] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. J. Wiley New York, 2001.
- [4] M. Ilyas, I. Mahgoub, and L. Kelly. *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*. CRC Press, Inc. Boca Raton, FL, USA, 2004.
- [5] I.T. Jolliffe. *Principal Component Analysis*. Springer, 2002.
- [6] S. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong. TAG: a Tiny AGgregation Service for Ad-Hoc Sensor Networks. In *Proceedings of the 5th ACM Symposium on Operating System Design and Implementation (OSDI)*, volume 36, pages 131 – 146. ACM Press, 2002.
- [7] S.R. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong. TinyDB: an acquisitional query processing system for sensor networks. *ACM Transactions on Database Systems (TODS)*, 30(1):122–173, 2005.
- [8] K.V. Mardia, J.T. Kent, J.M. Bibby, et al. *Multivariate analysis*. Academic Press New York, 1979.
- [9] Y. Yao and J. Gehrke. The cougar approach to in-network query processing in sensor networks. *ACM SIGMOD Record*, 31(3):9–18, 2002.



Index

- basis, 10, 24, 38, 52, 66, 80, 94, 108, 122, 136, 150
 - ordered, 10, 24, 38, 52, 66, 80, 94, 108, 122, 136, 150
- linear combination, 9, 23, 37, 51, 65, 79, 93, 107, 121, 135, 149
 - nontrivial, 9, 23, 37, 51, 65, 79, 93, 107, 121, 135, 149
 - trivial, 9, 23, 37, 51, 65, 79, 93, 107, 121, 135, 149
- linearly independent, 10, 24, 38, 52, 66, 80, 94, 108, 122, 136, 150
- maximal, 10, 24, 38, 52, 66, 80, 94, 108, 122, 136, 150
- matrix
 - scalar multiple of, 12, 26, 40, 54, 68, 82, 96, 110, 124, 138, 152
 - scalar multiplication, 12, 26, 40, 54, 68, 82, 96, 110, 124, 138, 152
- spanning set, 9, 23, 37, 51, 65, 79, 93, 107, 121, 135, 149
 - minimal, 10, 24, 38, 52, 66, 80, 94, 108, 122, 136, 150
- subspace, 8, 22, 36, 50, 64, 78, 92, 106, 120, 134, 148
- vector space
 - subspace of, 8, 22, 36, 50, 64, 78, 92, 106, 120, 134, 148

