

## まえがき

この本は、機械学習の技術を使って異常検知をやってみたい人のためのガイドブックとして書かれました。データに基づいて、なにか客観的な基準で変化の兆候をとらえて、いち早く意思決定をしてつぎの手を打つこと。これはあらゆるビジネスの基本ですが、異常検知はそのための第一歩となる大切な技術です。

筆者はこれまで、異常検知について開発した新たな手法を国際学会などで発表する傍ら、幅広い応用事例を手掛けてきました。その中で気づいたことが大きく二つあります。一つは、異常検知を実際にやってみたいと思っても、直接参考になる本を探すのが難しいということです。この研究を始めた10年ほど前、異常検知は統計学や制御工学の小さな一分野という扱いで、しかもその内容は、確率分布に関する理論的な解説がほとんどで、実データの荒々しさの前には使い物にならないというのが正直な感想でした。

気づいたことのもう一つは、異常検知においては、「機械学習」という比較的新しい学問分野で開発された技術が実用上とても役に立つということです。機械学習はもともと人工知能研究の一分野で、「機械が自動で学習する」という語感のとおり、大量のデータがあるときに、その中からパターンを見出すことを目標とする汎用技術です。機械学習は、2000年前後から理論・応用面で目覚ましい進歩を遂げました。どちらかといえば研究者側の発想で語られることの多いこの新しい技術分野を、異常検知という、実用ど真ん中の視点から語ってみたいと思います。

多くの場合、異常検知は、多大な経済的損失の防波堤としての重大な役割もっています。人の命にすら関わる場合もあります。そのような重要な問題に関わるエンジニアとしては、使う手法について完全な理解をもちたいと願うのが人情だと思います。そのような要請に応えるため、本書では天下り式の説明

をできるだけ排除し、理系の学部程度の統計学や線形代数の基礎知識さえあれば、本書のみで手法についての理解が得られるように努力しました。正規分布やカイ二乗分布といった主要な分布の分布形の導出さえ行っているのがその例です。ただし、細部の導出にこだわることで全体像が見えにくくなる危険も考えて、理論の詳細と考えられるところの節には「\*」印を付けてあります。本書をばらばらと眺めていただければ、日本の製造業の現場で普及しているホテリング  $T^2$  理論やマハラノビス=タグチ法はもちろん、機械学習の技術を使った基本的な異常検知の技術が、自己完結的に説明されていること、数式の書きっ放しではなく、実戦に使えるよう「手順」という形で内容をまとめていることなど、本書の特徴がおわかりいただけるかと思います。

本書は、電子情報通信学会「情報論的学習理論と機械学習研究会」が2012年に企画したチュートリアルのために準備した内容が基になっています。その機会を与えて下さり、またこの分野で研究を続けるきっかけをつくってくださった東京大学 山西健司 教授にはこの場を借りて御礼申し上げます。また、本書の執筆に際し、東京大学先端科学技術研究センターの矢入健久 博士には、その企画から細かい内容の確認に至るまで、ほとんど共著といえるほど多大なるご支援をいただきました。IBM 東京基礎研究所の勝木孝行氏には本書草稿について貴重かつ詳細な技術的コメントを得ました。ご支援に心から感謝いたします。

2015年1月

井手 剛

<http://ide-research.net/>

本書のサポートページを下記の URL に公開しています。

<http://ide-research.net/book/support.html>

章末問題の解答と誤植情報を掲載する予定です。

# 目 次

## 1. 異常検知の基本的な考え方

1.1	例題：健康診断	1
1.2	計算機に判定規則をつくらせたい	2
1.3	「確率分布」で正常パターンをつかむ	3
1.4	機械学習で確率分布を求める	6
1.5	やりたいことを具体的に整理する	7
1.6	異常の度合いを数値で表す	9
1.7	いろいろな手法を試してみる	10

## 2. 正規分布に従うデータからの異常検知

2.1	異常検知手順の流れ	15
2.2	1変数正規分布に基づく異常検知	17
2.2.1	ステップ1：分布推定	17
2.2.2	ステップ2：異常度の定義	19
2.2.3	ステップ3：閾値の設定	20
2.2.4	Rでの実行例	23
2.3	1変数のホテリング理論の詳細*	26
2.3.1	1変数正規分布の最尤推定	27
2.3.2	正規変数の和の確率分布（1次元）	28
2.3.3	標本分散の確率分布（1次元）	30

2.3.4	ホテリング統計量の確率分布 (1次元) .....	35
2.4	多変量正規分布に基づく異常検知 .....	37
2.4.1	ステップ1: 多次元正規分布の最尤推定 .....	37
2.4.2	ステップ2: 異常度の定義 .....	39
2.4.3	ステップ3: 閾値の設定 (ホテリングの $T^2$ 理論) .....	40
2.4.4	Rでの実行例 .....	42
2.5	多変数のホテリング理論の詳細* .....	44
2.5.1	多変量正規変数の和の分布 .....	44
2.5.2	多変量正規変数の平方和の分布 .....	45
2.5.3	ホテリング統計量の分布 .....	47
2.6	マハラノビス=タグチ法 .....	49
2.6.1	手法の概要 .....	49
2.6.2	Rでの実行例 .....	50
2.6.3	QR分解* .....	52
2.7	t分布による異常判定* .....	55
2.8	ホテリング理論の課題 .....	57
	章末問題 .....	58

### 3. 非正規データからの異常検知

3.1	分布が左右対称でない場合 .....	59
3.1.1	ガンマ分布の当てはめ .....	59
3.1.2	Rでの実行例 .....	62
3.1.3	カイ二乗分布による異常度の当てはめ .....	64
3.2	訓練データに異常標本が混ざっている場合 .....	65
3.2.1	正規分布の線形結合のモデル .....	65
3.2.2	期待値-最大化法: 期待値ステップ .....	67

3.2.3	期待値-最大化法：最大化ステップ	69
3.2.4	Rでの実行例	70
3.3	分布がひと山にならない場合：近傍距離に基づく方法	72
3.3.1	$k$ 近 傍 法	72
3.3.2	局所外れ値度	75
3.3.3	カーネル密度推定	78
3.3.4	Rでの実行例	82
3.4	分布がひと山にならない場合：クラスタリングに基づく方法	84
3.4.1	$k$ 平 均 法	84
3.4.2	混合正規分布モデル	86
3.4.3	異常度の定義とRによる実行例	90
3.5	期待値-最大化法の詳細*	93
3.5.1	イエンセンの不等式	94
3.5.2	最大化ステップ	97
3.6	支持ベクトルデータ記述法に基づく異常判定	98
3.6.1	データを囲む最小の球	99
3.6.2	Rでの実行例	100
	章 末 問 題	103

## 4. 性能評価の方法

4.1	基本的な考え方	105
4.2	正常標本精度と異常標本精度	106
4.2.1	正常標本に対する指標	106
4.2.2	異常標本に対する指標	107
4.3	異常検出能力の総合的な指標	108
4.3.1	分岐点精度とF値	109

4.3.2	ROC 曲線の下部面積	112
4.4	モデルのよさの検証	113
4.4.1	モデル選択問題	113
4.4.2	交差確認法	114
4.4.3	赤池情報量規準とベイズ情報量規準	116
4.4.4	赤池情報量規準と平均対数尤度*	117
4.4.5	ベイズ情報量規準と周辺尤度*	119
	章末問題	121

## 5. 不要な次元を含むデータからの異常検知

5.1	次元削減による異常検知の考え方	123
5.2	主成分分析による正常部分空間の算出	124
5.2.1	分散最大化規準による正常部分空間	125
5.2.2	ノルム最大化規準による正常部分空間	127
5.2.3	二つの規準の等価性と特異値分解	129
5.3	主成分分析による異常検知	132
5.3.1	異常度の定義	132
5.3.2	ホテリングの $T^2$ との関係	134
5.3.3	次元 $m$ の選択	135
5.3.4	Rでの実行例	136
5.4	確率的な主成分分析による異常検知*	140
5.4.1	主成分分析の確率的モデル	140
5.4.2	平均ベクトルの推定	141
5.4.3	確率的な主成分分析の期待値-最大化法	143
5.4.4	$\sigma^2 \rightarrow 0$ の極限と次元数 $m$ の決定	145
5.4.5	確率的な主成分分析による異常度の定義	146

5.5	カーネル主成分分析による異常検知*	147
5.5.1	正常部分空間の算出	148
5.5.2	Rでの実行例	150
5.5.3	異常度の定義 ( $m = 1$ )	152
5.5.4	異常度の定義 ( $m > 1$ )	155
章 末 問 題		158

## 6. 入力と出力があるデータからの異常検知

6.1	入出力がある場合の異常検知の考え方	159
6.2	線形回帰モデルによる異常検知	161
6.2.1	問題の定義	161
6.2.2	最小二乗法としての最尤推定	162
6.2.3	異常度の定義	164
6.3	リッジ回帰モデルと異常検知	165
6.3.1	リッジ回帰の解	165
6.3.2	定数 $\lambda$ の決定	167
6.3.3	異常度の定義	169
6.3.4	Rでの実行例	170
6.4	偏最小二乗法と統計的プロセス制御 (1次元出力)	173
6.4.1	問題の設定	173
6.4.2	正規直交基底による回帰モデルの変換	174
6.4.3	NIPALS法 (1次元出力)	176
6.4.4	異常度の定義と異常検知手順	178
6.5	正準相関分析による異常検知	179
6.5.1	問題設定	180
6.5.2	一般化固有値問題としての正準相関分析	181

6.5.3	特異値分解による解と異常検知	183
6.6	ベイズ的線形回帰モデルと異常検知*	185
6.6.1	最大事後確率解としてのリッジ回帰	186
6.6.2	パラメター $\sigma^2$ の決定	188
6.6.3	異常度の定義	190
章 末 問 題		192

## 7. 時系列データの異常検知

7.1	近傍法による異常部位検出	194
7.1.1	スライド窓による時系列データの変換	194
7.1.2	異常部位検出問題	196
7.1.3	Rでの実行例	198
7.2	特異スペクトル変換法	199
7.2.1	特徴的なパターンの算出	199
7.2.2	変化度の定義	202
7.2.3	Rでの実行例	203
7.3	自己回帰モデルによる異常検知	205
7.3.1	1変数の自己回帰モデル	206
7.3.2	ベクトル自己回帰モデル	207
7.3.3	次数 $r$ の決定	209
7.3.4	異常度の定義と R での実行例	210
7.4	状態空間モデルによる異常検知*	212
7.4.1	線形状態空間モデル	213
7.4.2	部分空間同定法：状態系列の推定	214
7.4.3	部分空間同定法：未知パラメター A, C, Q, R の推定	219
7.4.4	状態系列の逐次推定法：カルマンフィルタ	221

7.4.5 状態空間モデルを用いた異常検知	224
章 末 問 題	227

## 8. よくある悩みとその対処法

8.1 数式を使いたくありません	229
8.2 モデルが変わってゆくのですが	230
8.3 変数の値の範囲が変で困っているのですが	232
8.3.1 ロジスティック変換	233
8.3.2 ボックス=コックス変換	234
8.4 正規分布の結果がおかしいのですが	235
8.5 データがベクトルになっていないのですが	237
8.6 異常の原因を診断したいのですが	238
8.7 分類問題にははいけませんか	239
8.8 さらに発展的な知識を得るために	240

## 付 録 242

A.1 有用な R のパッケージ	242
A.2 確率変数の変換	244
A.2.1 確率密度関数と周辺化	244
A.2.2 条件付き確率と独立性	245
A.2.3 逆変換が一意に定義できる場合の変換	246
A.2.4 $M$ 変数から 1 変数への変換	247
A.3 有用な行列公式	249
A.3.1 ブロック行列と逆行列	249
A.3.2 行 列 式	251
A.3.3 行 列 の 微 分	252

A.4	正規分布の性質のまとめ	253
A.4.1	正規分布の確率密度関数の導出	253
A.4.2	単位球の表面積	255
A.4.3	正規変数の和の分布	256
A.4.4	多変数正規分布の分割公式	257
A.4.5	多変数正規分布とベイズ公式	259
A.5	カーネル密度推定における平均積分二乗誤差の漸近形	260
A.6	等式制約付きの非線形最適化	262
引用・参考文献		266
索引		269
Rコマンド索引		274

# 1

## 異常検知の基本的な考え方

この章では、簡単な例題を基に、異常検知の問題設定と、基本となる考え方を眺めてゆきます。

### 1.1 例題：健康診断

あなたがある高校の校医さんで、1 学年 200 人の身長と体重、さらに問診結果のデータを一覧できる立場にあると考えましょう。そして、このデータから「普通と変わった振舞い」を見つけ出す規則をつくることを考えましょう。常識的に頭に浮かぶのは、例えばこういう規則です。

「体重が 80 kg 以上なら太りすぎ」

なぜこう思うかといえば、頭の中に「普通の人ならこれくらいだ」というイメージがあり、それから外れる人を肥満とみなしているからです。

しかし、身長が高い人は当然体重も多いでしょうから、単に体重だけで太りすぎを判定するのは筋が通らない気がします。そこで世界保健機構（WHO）では、つぎのような指標を定め、BMI（body mass index）と呼んでいます。

$$\text{BMI} = (\text{体重 [kg]}) \div (\text{身長 [m] の 2 乗})$$

そうして、これが 25 を超えると肥満、すなわち「普通と変わった振舞い」にあると定義しています<sup>†</sup>。なぜ WHO が 25 以上を肥満だと考えているかといえば、正常で健康な人が大体 BMI 25 未満に収まっているという臨床データを

---

<sup>†</sup> 体重は身体の体積に比例するはずなので、一見身長<sup>3</sup>のほうの方が自然な気もしますが、赤ちゃんは別にして、大人の場合は体重と体積は比例しないそうです。

## 2 1. 異常検知の基本的な考え方

もっているからです。ただ、WHOの基準にはいろいろな国のいろいろな人種が混ざっているはずですから、高校生の健康のよい指標にはならないかもしれません。この場合、問診のデータを基に、「BMIがこのくらいなら不健康な人が多い」というような区切りの値（これを閾値（しきいち）と呼びます）をいろいろ調整して定めることになるのだと思います。

これはとても簡単な例ですが、異常検知の考え方をよく説明する例になっています。つまり異常検知の基本は、正常となるモデル（「普通の人ならBMIはこのくらいだ」）をデータからつくり、そのモデルから外れるものを異常とすることです。つぎの節で、もう少し詳しくこの問題を定義して、異常検知の問題設定についての理解を深めましょう。

### 1.2 計算機に判定規則をつくらせたい

まず、一般的に異常検知を行うにはなにが必要か考えてみましょう。例えば、上の健康診断の場合、「健康な人であれば、BMIの値がこれくらいの範囲におさまるはず」というような判定基準が必要になります。この判定基準をより一般的な言葉で表せば、正常と異常を区別するためのなんらかの「知識」ということになります。つまり、異常検知には「知識」が必要です。

「機械学習で異常検知を行う」というのは結局、この「知識」を（人間が事前に用意するのではなく）機械学習の手法を用いて過去のデータから計算機に見つけ出させるということを意味しています。この流れを模式的に図1.1に書きました。知識は統計的モデルの形で表現され、統計モデルのパラメーターや係

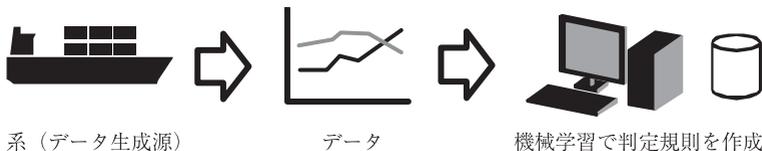


図 1.1 機械学習を用いた異常検知の流れ（データに基づいて統計的モデルの形の知識が自動で学習される）

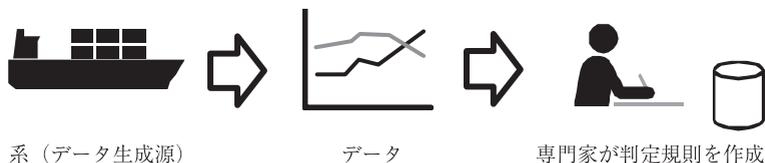


図 1.2 伝統的な異常検知の流れ（知識の生成源は主に人間で、典型的には人間が可読な IF-THEN ルールの形で知識が蓄積される）

数など、一般にはなにかの数値の集まりとなります。それらの数値を「読む」ためには一般には統計的モデルへの理解が必要ですが、統計モデルは、例えば「異常度」という形で現実世界に対し有用な知見を与えるように設計されます。

統計的機械学習の手法が発達する以前は、知識の生成源は人間でしかありえませんでした。知識は基本的に人間が読める形での IF-THEN ルールで記述されることが想定され、このルールを大量にためておけば最高の専門家の思考過程が再現できると考えられてきました。いわゆるエキスパートシステムです（図 1.2）。IF-THEN ルールは可読性の観点で魅力的ですが、人工知能の長い研究史が教えるところによれば、人間が明示的に列挙できるルールの多様性は、現実世界の多様性に比べて桁違いに乏しく、人間を主たる判定ルール生成源として監視システムを構築するアプローチはほぼ必ず失敗することがわかっています。実用的な監視システムの構築には、データからパターンを自動学習する機能が必要です。

最近、計測・データ保存の技術が発展して膨大なデータが集められるようになってきています。その一方で、適切な異常検知ルールやモデルを作成できる熟練エンジニアが減りつつあるといわれています。従来のエキスパートシステムが抱えていた本質的な課題（しばしば知識獲得のボトルネックと呼ばれます）を、機械学習という新しい解析技術を使って乗り越える機運が高まってきたといえるでしょう。本書はそのためのガイドブックになることを意図しています。

### 1.3 「確率分布」で正常パターンをつかむ

冒頭の健康診断の例では、身長と体重という二つの値を考えました。これを

#### 4 1. 異常検知の基本的な考え方

一般に**変数**と呼ぶことにします。生徒によって値がいろいろと変わるからです。いまの場合は、二つ変数がありますが、これを「入力変数は2次元である」ということにします。2次元以上でも同様です。2次元以上の変数をまとめて、太字で  $\mathbf{x}$  というように表します。いまの場合、生徒の名前を仮に一郎さん、二郎さん、三郎さん、…としておくと、 $i$  郎さんについての変数を

$$\mathbf{x}^{(i)} = \begin{pmatrix} x_{\text{身長}}^{(i)} \\ x_{\text{体重}}^{(i)} \end{pmatrix} \quad (1.1)$$

のように定義できます。 $x_{\text{身長}}^{(i)}$  は、 $i$  郎さんの身長を表す記号で、例えば二郎さんについては  $x_{\text{身長}}^{(2)} = 178 \text{ cm}$  というような感じです。同様に、二郎さんの体重を  $x_{\text{体重}}^{(2)} = 63 \text{ kg}$  のように表します。以下、複数の変数をまとめて列ベクトルとして表します<sup>†</sup>。この二郎さんの例だと、 $\mathbf{x}^{(2)} = \begin{pmatrix} 178 \\ 63 \end{pmatrix}$  という感じです。

データには、身長・体重に加えて問診の結果もあると想定しました。話を簡単にするため、「不健康度」が記録されているとしましょう。例えば完全に健康なら0、まったく病的なら1、というような値を定義できるかもしれません。これはいってみれば、それぞれの人の日ごろの節制を表す結果、あるいは出力というべき変数です。これをやはり  $i$  郎さんに対して  $y^{(i)}$  という記号で表すことにしましょう。これらの記号を使うと、与えられたデータ  $\mathcal{D}$  は、全生徒の数  $N = 200$  に対して

$$\mathcal{D} = \{(\mathbf{x}^{(1)}, y^{(1)}), (\mathbf{x}^{(2)}, y^{(2)}), \dots, (\mathbf{x}^{(i)}, y^{(i)}), \dots, (\mathbf{x}^{(N)}, y^{(N)})\}$$

というように書けます。 $\{\cdot\}$  は集合を表す記号で、 $(\cdot)$  は「組」を表す記号という意味で使っています。集合の場合、要素の順番はいつでもよいですが、組の場合は要素の順番には意味があることに注意しましょう。

さて、これらの準備の下で、冒頭で紹介した異常検知の基本的考え方をより

<sup>†</sup> 「列」ベクトルとするのはただの習慣で深い意味はありませんが、行列とベクトルの積の表記がきれいになるというメリットがあり、線形代数での標準的記法です。

「それらしく」表現してみましょう。異常検知モデルの構築は一般に3ステップあります。

- 1) ステップ1 (分布推定): 正常のモデルをつくる。通常、「正常」といってもある範囲でばらつくものですから、先の例ですと、「標準体型」の人の、身長と体重に関するばらつきを含んだモデル（確率分布）をつくることになります。
- 2) ステップ2 (異常度の定義): 正常からのずれの度合い、すなわち異常度を定義する。この場合、すごく体重が重い人とか、身長と体重のバランスがひどく崩れている人が高い異常度をもつように異常度を設計することになります。
- 3) ステップ3 (閾値の設定): 異常度がある値より大きいと異常、と判定できるような区切りの値（閾値）をデータから求めます。いまの場合、正常人のモデルと問診結果  $y$  を照合しつつその範囲を決めることになります。

分布、異常度、閾値を、異常検知の三大要素と呼んでよいでしょう。機械学習を用いて異常検知を行うということは、これらの要素を、機械学習を使ってデータから最適に決めることです。

なお、実応用上、これらの三要素が一部融合した形で扱われることがあります。特に、ステップ1と2の分布推定と異常度の定義を融合させ、全体を2ステップで取り扱う場合がしばしばあります。しかし本書では、異常検知技術全体に俯瞰的な視野を提示するという観点から、確率分布の存在を基本に考えます。

では、「確率分布を求める」というのはどういうことでしょうか。通常、正常のモデルをつくるためには、つぎの二つの作業を行うのが普通です。

- (1) 未知のパラメーターを含む確率分布モデルを仮定する。
- (2) 未知パラメーターをデータに合わせこむ。

これは「データからのモデルの学習」、すなわち、機械学習そのものです。例えば健康診断の例で、正規分布を使ったとすると、これは、身長と体重について  $N = 200$  人のデータから平均と分散を計算することに他なりません。これ

は誰でも知っている演算ではありますが、まさに「データからのモデルの学習」であり、機械学習で目標にすることそのものです。本書を手にとった読者の中には、機械学習という分野についてまったく予備知識がない人もいるかもしれません。でも大丈夫。機械学習ですることは、基本的に単純、正規分布の平均・分散を計算しているのと同じようなことです。次節でもう少し詳しく例を挙げます。

### 1.4 機械学習で確率分布を求める

正規分布の平均と分散の例は、機械学習でやることの最も単純な例でしたが、もう少し細かく問題を分けることができます。データから正常モデルを構築するには、対象となる系の性質に応じて、つぎのような問題を解く必要があります。

- (1) **密度推定問題** 入出力に区別がなく、また、データの観測順序が重要でない場合。先の例だと、体重と身長に関する標準的なばらつきを表す確率分布  $p(\mathbf{x})$  を、データ  $\mathcal{D}$  から求めることに当たります。2章と3章で主に扱います。
- (2) **次元削減問題** これも  $p(\mathbf{x})$  を  $\mathcal{D}$  から求める問題ですが、 $\mathbf{x}$  が多次元で、しかも、データのパターンの把握に役に立つ次元とそうでない次元が混在している状況を想定します。このような場合、役に立たない次元を削減し、役に立つ次元だけでモデルをつくるのが理想的です。そのための手法を5章で扱います。
- (3) **回帰問題** ある入力に対する出力に興味があるが、観測データの順序が重要でない場合。先の例だと、身長と体重のデータ  $\mathbf{x}$  を与えたときの、不健康度  $y$  の値についての条件付き確率分布  $p(y | \mathbf{x})$  を  $\mathcal{D}$  から求めることに当たります<sup>†</sup>。6章で主に扱います。
- (4) **分類問題** 上と同じく、 $p(y | \mathbf{x})$  を  $\mathcal{D}$  から求める問題ですが、この

---

<sup>†</sup>  $p(y | \mathbf{x})$  の “|” は条件付き確率を表す記法。付録 A.2.2 項に基礎事項をまとめておきます。

場合、 $y$  は、例えば、0 (健康) と 1 (不健康) の 2 種類の値、というような離散値をとるのが特徴です。実用上は、健康か否かの結果だけでなく、その度合いを知りたいことが多いので、本書では分類問題としての定式化は主たる興味の対象とはしませんが、3.3 節の近傍法の文脈で、また 8.7 節で一般的な注意をまとめています。

- (5) **時系列解析問題** これは、観測データの順番に意味がある場合、特に典型的なのはデータが時系列になっている場合です。例えば、一郎さんの身長と体重を 5 年間ずっと追跡する、というような状況です。この場合、ある時刻  $t$  における観測値のモデルが重要ですから、出力変数の有無に応じてそれぞれ  $p(\mathbf{x} | t)$  または  $p(y | \mathbf{x}, t)$  を  $\mathcal{D}$  から求めることになります。7 章で扱います。

機械学習を使って異常検知をするとは、これらの問題を機械学習の分野で発展してきた手法を駆使して解き、異常度と閾値を問題に応じた適切な形で決めることを意味します。

## 1.5 やりたいことを具体的に整理する

上に述べた分類は機械学習側から述べたものですが、これを実問題側から見てみましょう。本書では例えばつぎのような問題パターンを扱っています。

- (1) 手元にあるデータ  $\mathcal{D}$  の中に、測定エラーによる値が混じっているかもしれないので、そういう値を取り除きたい。これはデータクレンジング (データ洗浄) と呼ばれる問題です。例えば 2 章で述べる Davis データの中には、身長と体重を取り違えた標本が含まれています。本書の多くの箇所でも例題として紹介されます。
- (2) 過去に観測したデータ  $\mathcal{D}$  を基にして、いま得られた観測値  $\mathbf{x}$  の異常度を計算したい。過去 5 年間の健診データを基に、健康人のモデルをつくり、個々人に対する最新のデータに対して健康判定を行うイメージです。通常、外れ値検出の問題、すなわち、正常時に期待される値から外れた

## 8 1. 異常検知の基本的な考え方

かどうかを判定する問題となります。2章のホテリング理論を手始めに、本書の多くの箇所で紹介されます。なお、 $\mathcal{D}$ の中の標本に外れ値検出問題を解くのがデータクレンジングです。

- (3) データ  $\mathcal{D}$  の中の標本を、いくつかの塊に束ねて、データの中のパターンを把握したい。これはクラスタリングです。健診のデータだと、おそらく、男子と女子で異なる傾向があるでしょうから、二つのクラスターが現れるかもしれません。3.4節で別の例が挙げられます。
- (4) 入力  $\boldsymbol{x}$  と出力  $y$  がある決まった関係にあると想定されるときに、その関係の崩れを検知したい。これは前節でも言及した回帰問題として解けます。また、入出力が複数あるときには、**正準相関分析**も有用です。6章で詳しく説明されます。
- (5) 時系列データを眺めて、他と違った振舞いをする部分を同定したい。**異常部位検出**と呼ばれる問題です。例えば、ある人の心電図データを眺めて異常な拍動を呈する部位を抜き出すことは、診断に有用だと思います。図 7.1 が実例です。
- (6) 時系列データを眺めて、なにやら傾向が変わった時点で警報を出す。これは**変化点検知**という問題です。これも図 7.1 に実例があります。
- (7) 高次元のデータのばらつきを可視化して、大雑把に状況を把握したい。これは**主成分分析**という手法を使えば可能です。実例として図 5.5 は、自動車の異なる車種の属性データの分布を描いた例です。
- (8) 二つ（またはそれ以上）のデータセットがあるとき、違いがあるのかわるのかを調べたい。東京の生徒と北海道の生徒の体格に違いがあるのだろうか、のような問題で、これは**変化解析問題**と本書で呼ぶ問題です。図 5.9 に示すように、自動車なり船舶なりの観測データを基に不具合を解析したい、という問題でも現れます。

半導体の生産監視の分野では、FDC<sup>†</sup>という名前で異常検知のツールが実用化され販売されていますが、これは外れ値検出問題を解いていることに対応し

<sup>†</sup> Fault Detection and Classification の略。

# 索引

## 【あ】

アイビス (IBIS) 241  
赤池情報量規準 (Akaike information criterion) 116, 209  
アップサンプリング (up-sampling) 239

## 【い】

イエンセンの不等式 (Jensen's inequality) 94, 143  
異常検知の三大要素 (three key factors of anomaly detection) 5  
異常度 (anomaly score) 5, 9, 16  
異常標本精度 (anomalous sample accuracy) 107  
異常部位検出 (discord discovery) 8, 194, 196  
異常網羅率 (coverage) 108  
1クラスサポートベクトルマシン (one-class support vector machine) 99  
1クラス支持ベクトル分類器 (one-class support vector machine) 99  
1次関数 (linear function) 161  
一般化逆行列 (generalized inverse matrix) 236  
一般化交差確認法 (generalized cross validation) 168  
一般化固有値方程式 (generalized eigenvalue equation) 183, 218  
 $\epsilon$  近傍法 ( $\epsilon$ -neighbor method, method of  $\epsilon$ -neighbors) 74  
異方性 (anisotropy, heterogeneity) 39

## 【う】

ウィシャート分布 (Wishart distribution) 41, 47  
上に凸 (convex upward, concave) 94  
ウッドベリー行列恒等式 (Woodbury matrix identity) 168, 189, 222, 250

## 【え】

エッカート=ヤングの定理 (Eckart-Young theorem) 158  
エビデンス (evidence) 188  
エビデンス近似 (evidence approximation) 188  
エルボー則 (elbow rule) 136  
エントロピー (entropy) 253

## 【お】

凹関数 (concave function) 94  
重みづけ (weighting) 239  
オンライン更新則 (online updating algorithm) 232

## 【か】

カーネル関数 (kernel function) 78, 149  
カーネル主成分分析 (kernel principal component analysis) 132, 149  
カーネルトリック (kernel trick) 100, 150, 191  
カーネル法 (kernel method) 100, 147  
カーネル密度推定 (kernel density estimation) 78, 116, 243, 260  
回帰 (regression) 6, 160  
階数 (rank) 129, 166

階数1更新 (rank-one update) 228  
カイ二乗分布 (Chi-squared distribution) 21  
ガウス過程回帰 (Gaussian process regression) 243  
ガウス積分 (Gauss integral) 254  
可観測行列 (observability matrix) 216  
可観測性行列 (observability matrix) 216  
核関数 (kernel function) 78, 82, 260  
確認データ (validation data) 26, 105  
確率変数 (random variable) 244  
確率密度関数 (probability density function) 18, 247, 257  
隠れ変数 (hidden variable) 93, 102  
過剰適合 (overfitting) 114  
滑走窓 (sliding window) 196  
カルマンゲイン (Kalman gain) 224  
カルマンフィルタ (Kalman filter) 224  
カルマン利得行列 (Kalman gain) 224  
観測変数 (observable variable) 213  
ガンマ関数 (gamma function) 21, 60, 255  
ガンマ分布 (gamma distribution) 59

## 【き】

規格化条件 (normalization condition) 244  
帰属度 (membership probability) 69, 90, 237

期待値-最大化 (expectation-maximization) 法  
67, 68, 89, 93, 237  
逆行列 (inverse matrix)  
52, 134,  
146, 164, 166, 189, 249  
逆変換 (inverse transformation)  
246  
球座標 (spherical coordinates) 30, 46, 79, 255  
偽陽性率 (false positive rate)  
107  
共分散行列 (covariance matrix) 37, 42, 86, 181, 208  
行列式 (determinant)  
34, 37, 236, 247, 251  
——の補題 (matrix determinant lemma) 251  
行列 2 ノルム (matrix 2-norm) 155  
行列ノルム (matrix norm) 155  
行列  $p$  ノルム (matrix  $p$ -norm) 156  
局所外れ値度 (local outlier factor) 76  
曲線下部面積 (area under curve) 112  
近傍有効距離 (reachability distance) 76

【く】  
組 (tuple) 4  
クラスタリング (clustering)  
8, 84, 87, 88, 103, 243  
グラスマン多様体 (Grassmann manifold) 157  
グラム行列 (Gram matrix)  
128  
グラム=シュミットの直交化法 (Gram-Schmidt orthogonalization) 53  
クリロフ部分空間 (Krylov subspace) 205  
クロネッカーのデルタ (Kronecker's delta) 248  
訓練データ (training data)  
26, 65

## 【け】

計画行列 (design matrix)  
162  
経験分布 (empirical distribution)  
73, 103, 114, 119, 232  
経験ベイズ法 (empirical Bayes method) 188  
結合分布 (joint distribution)  
245  
決定木 (decision tree) 243  
検証データ (validation data)  
26, 198

## 【こ】

交差確認 (cross validation)  
114, 209  
5 重交差確認 (5-fold cross validation) 法 114  
誤報率 (false alarm rate)  
107  
固有値 (eigenvalue)  
33, 126, 135  
固有値方程式 (eigenequation) 33, 131, 183  
固有値問題 (eigenvalue problem) 33, 129, 145  
固有ベクトル (eigenvector)  
33, 126, 131, 156, 236  
コレスキー分解 (Cholesky factorization, Cholesky decomposition) 55, 236  
混合ガウス分布 (Gaussian mixture distribution, mixture of Gaussian distributions) 66  
混合正規分布 (normal mixture distribution, mixture of normal distributions) 66, 87, 117

## 【さ】

再構成誤差 (reconstruction error) 133  
最小二乗法 (least squares)  
162

最大事後確率 (maximum a posteriori) 167, 187  
最尤推定 (maximum likelihood estimation)  
18, 27, 37, 60, 93,  
116, 141, 162, 208, 219  
差込み法 (plug-in method)  
81  
サポートベクトルマシン (support vector machine)  
239, 243  
散布行列 (scatter matrix)  
125

## 【し】

閾値 (threshold)  
2, 23  
シグモイド関数 (sigmoid function) 233  
時系列解析 (time-series analysis) 7, 240, 243  
次元 (dimension) 4  
次元削減 (dimensionality reduction)  
6, 127, 175  
自己一致 (self-match) 196  
自己回帰モデル (autoregressive model) 206  
自己組織化マップ (self-organizing map) 242  
事後分布 (posterior distribution) 186  
支持ベクトルデータ記述法 (support vector data description) 100, 238  
支持ベクトル分類器 (support vector machine) 243  
次数 (order) 206, 208  
システム同定 (system identification) 241  
事前分布 (prior distribution)  
121, 141, 157, 186  
自明な一致 (trivial match)  
196  
尺度変換 (scale transformation) 233  
シュア補行列 (Schur's complement) 250  
集合 (set) 4

自由度 (degrees of freedom) 21, 32, 36, 41, 56, 64  
 周辺化 (marginalization) 87, 102, 121, 188, 190, 245  
 周辺化尤度 (marginalized likelihood) 121, 143  
 周辺分布 (marginal distribution) 245  
 周辺尤度 (marginal likelihood) 121, 188  
 主軸問題 (principal axis problem) 33, 47, 48  
 受信者操作特性曲線 (receiver operating characteristic curve) 112  
 主成分 (principal component) 131  
 主成分分析 (principal component analysis) 8, 124, 132, 140, 243  
 シュテューフェル多様体 (Stiefel manifold) 46  
 主部分空間 (principal subspace) 124  
 条件付き確率 (conditional probability) 6, 16, 86, 245  
 条件付き分布 (conditional distribution) 245  
 情報量 (information) 10, 255  
 情報論的学習理論ワークショップ (Information-based induction sciences workshop) 241  
 シルベスターの行列式補題 (Sylvester's determinant lemma) 34, 189, 251  
 心電図データ (electrocardiogram data) 195  
 真陽性率 (true positive ratio) 108

**【す】**

スケール因子 (scale factor) 21, 31, 41, 60  
 スライド窓 (sliding window) 196

**【せ】**

正規直交基底 (orthonormal basis) 125, 174  
 正規分布 (normal distribution) 10, 18, 37, 253  
 正規変数 (normal variable) 28, 30, 45, 47, 256  
 制御理論 (control theory) 241  
 正準相関係数 (canonical correlation coefficient) 183  
 正準相関分析 (canonical correlation analysis) 8, 179  
 正準変数 (canonical variable) 184  
 正常標本精度 (normal sample accuracy) 106  
 正常部分空間 (normal subspace) 124, 149  
 正則化項 (regularization term, regularizer) 166  
 正答率 (detection rate) 107  
 性能分岐点 (break-even point) 109  
 跡 (trace) 38  
 積分消去 (integrate out) 245  
 積分二乗誤差 (integrated squared error) 79  
 線形関数 (linear function) 161  
 線形状態空間モデル (linear state-space model) 213  
 潜在変数 (latent variable) 141

**【そ】**

相関規則学習法 (association rule mining) 226  
 相互部分空間法 (mutual subspace method) 157  
 ソフトセンサー (soft sensor) 180

**【た】**

対数変換 (log transformation) 234

対数尤度比 (log-likelihood ratio) 121  
 体積要素 (volume element) 244  
 第2種最尤推定 (type 2 maximum likelihood) 188  
 ダウンサンプリング (down-sampling) 239  
 多項式カーネル (polynomial kernel) 150  
 多次元のガンマ関数 (multivariate gamma function) 41  
 多重共線性 (multicollinearity, collinearity) 165, 173  
 単位球 (unit sphere) 255  
 単位空間 (unit space) 50  
 単純ベイズ分類 (naïve Bayes classification) 243  
 単峰的 (unimodal) 74

**【ち】**

逐次更新則 (sequential updating algorithm) 232  
 知識獲得のボトルネック (knowledge acquisition bottleneck) 3  
 中心化行列 (centering matrix) 33, 47, 125  
 中心極限定理 (central limit theorem) 122  
 直交行列 (orthogonal matrix) 33, 46, 53, 130, 158  
 直交変換 (orthogonal transformation) 32, 34

**【て】**

テイラー展開 (Taylor expansion) 152, 234, 261, 264  
 データクレンジング (data cleansing) 7, 25, 115  
 データ洗浄 (data cleansing) 7  
 テストデータ (test data) 26, 110

デルタ関数 (delta function) 31, 73, 248  
 点推定 (point estimation) 162  
**【と】**  
 動径基底関数展開 (radial basis function expansion) 79  
 動径座標 (radial coordinate) 30  
 統計的機械学習 (statistical machine learning) 3, 88, 240  
 統計的独立性 (statistical independence) 246  
 統計のプロセス制御 (statistical process control) 173  
 同時分布 (simultaneous distribution) 32, 245, 255  
 動的時間伸縮法 (dynamic time-warping) 197  
 特異スペクトル解析 (singular spectrum analysis) 201  
 特異スペクトル変換 (singular spectrum transformation) 201  
 特異値分解 (singular value decomposition) 130, 183, 201  
 独立性 (independence) 246  
 トレース (trace) 38, 169, 252  
 トレンド抽出 (trend extraction) 241  
**【な～の】**  
 内部状態 (internal state) 213  
 二項分布 (binomial distribution) 233  
 二次形式の標準形 (canonical form of quadratic form) 33  
 ネイマン=ピアソンの補題 (Neyman-Pearson lemma) 121  
 ノルム最大化規準 (max norm criterion) 127

**【は】**

バースト性 (bursty nature) 234  
 パーセンタイル (percentile) 16, 62, 229  
 パーセント点 (percentile) 105  
 バイアス (bias) 119  
 外れ値検出 (outlier detection) 7, 15, 20, 137, 196, 199  
 ハット (hat) 18  
 汎関数 (functional) 253  
 バンド幅 (band width) 79, 80, 260

**【ひ】**

非正値二次形式 (non-positive quadrature) 260  
 左特異ベクトル (left singular vector) 131  
 ヒット率 (hit ratio) 108  
 一つ抜き交差確認法 (leave-one-out cross validation) 26, 80, 115, 167  
 標準化 (standardization) 174  
 標準正規分布 (standard normal distribution) 58, 81, 255  
 標準偏差 (standard deviation) 19  
 標本共分散行列 (sample covariance matrix) 124  
 標本分散 (sample variance) 18, 30, 32  
 標本平均 (sample mean) 18, 39  
 頻度主義 (frequentism) 157

**【ふ】**

ブートストラップサンプリング (bootstrap sampling) 239  
 不均衡データ (imbalanced data) 239  
 復元抽出 (sampling with replacement) 239

普通の最小二乗法 (ordinary least squares) 164, 175, 207  
 負の対数尤度 (negative log likelihood) 10, 16, 19, 52, 62, 72, 140, 164  
 部分空間同定法 (subspace identification method) 214  
 部分最小二乗法 (partial least squares) 173  
 部分時系列 (time series subsequence, subsequence) 196  
 部分時系列クラスタリング (time-series subsequence clustering) 226  
 不偏分散 (unbiased variance) 24  
 ブロック行列 (block submatrix) 249  
 フロベニウスノルム (Frobenius norm) 156  
 分位点 (quantile) 16, 20, 62, 64, 161, 171  
 分割公式 (partition formula) 257  
 分岐点精度 (break-even accuracy) 109  
 分散最大化規準 (max variance criterion) 125  
 分布推定 (distribution estimation) 5, 15  
 分類 (classification) 6, 88

**【へ】**

平均情報量 (average information) 10, 255  
 平均積分二乗誤差 (averaged integral squared loss) 261  
 平均対数尤度 (averaged log likelihood) 118  
 ベイジアンネットワーク (Bayesian network) 238  
 ベイズ主義 (Bayesianism) 157  
 ベイズ情報量規準 (Bayesian information criterion) 91, 117

ベイズの定理 (Bays' thorem) 87, 143, 186, 221, 246  
 平方完成 (completion of the squares) 257  
 ベータ関数 (beta function) 21  
 ベータ分布 (beta distribution) 233  
 べき乗法 (power method) 158  
 ベクトル自己回帰モデル (vector autoregressive model) 208  
 ベクトル 2 ノルム (vector 2-norm) 78  
 ベクトル量子化 (vector quantization) 242  
 ベクトル  $p$  ノルム (vector  $p$ -norm) 156  
 変化解析 (change analysis) 8, 154, 202, 238  
 変化点検知 (change-point detection) 8, 199  
 偏最小二乗法 (partial least squares) 173  
 変数 (variable) 4  
 変分ベイズ法 (variational Bayes method) 237

**【ほ】**

忘却率 (forgetting rate) 231  
 ボックス=コックス変換 (Box-Cox transformation) 234  
 ホテリング統計量 (Hotelling's statistic) 21, 41  
 ホテリングの  $T^2$  (Hotelling's T-squared statistic) 21, 41  
 ホテリング理論 (Hotelling's theory) 8, 15, 16, 20, 26, 37, 44, 57

**【ま】**

マハラノビス距離 (Mahalanobis' distance) 10, 39, 42, 48, 49, 52, 54, 64, 91, 165, 191, 236, 243

マハラノビス=タグチ=グラム=シユミット法 (Mahalanobis-Taguchi-Gram-Schmidt method) 53  
 マハラノビス=タグチ法 (Mahalanobis-Taguchi method) 49, 120  
 間引き (down-sampling, sampling) 239  
 マルコフ性 (Markov property) 213

**【み】**

右特異ベクトル (right singular vector) 131  
 水増し (up-sampling, data augmentation) 239  
 密度推定 (density estimation) 6

**【め】**

面積要素 (area element) 244  
 面素 (area element) 30, 46, 255

**【も】**

モーメント法 (method of moment) 61, 179  
 目的関数 (objective function) 60, 86  
 モデル選択 (model selection) 114, 188, 209  
 モンテカルロフィルタ (Monte Carlo filter) 241

**【や】**

ヤコビアン (Jacobian) 247

**【ゆ】**

ユークリッド距離 (Euclidean distance) 76, 78, 85, 197  
 ユークリッドノルム (Euclidean norm) 155  
 有効次元 (effective dimension) 65  
 尤度 (likelihood) 9, 27, 97

尤度関数 (likelihood function) 27, 162

**【よ】**

予測分布 (predictive distribution) 16, 161, 190, 225

**【ら】**

ラグ (lag) 200  
 ラグランジュ関数 (Lagrangian, Lagrange function) 264  
 ラグランジュ乗数 (Lagrange multiplier) 69, 96, 126, 127, 177, 182, 254, 264  
 ラグランジュ未定乗数法 (method of Lagrange multiplier) 264  
 ラッソ回帰 (lasso regression) 242  
 ラプラスカーネル (Laplace kernel) 150  
 ランク落ち (rank deficient) 235  
 ランダム森 (random forest) 243

**【り】**

リコール (recall) 108  
 リッジ回帰 (ridge regression) 166  
 リプレゼンター定理 (representer theorem) 191  
 履歴行列 (trajectory matrix) 200

**【る〜ろ】**

累積寄与率 (cumulative contribution ratio) 135  
 連続型確率変数 (continuous random variable) 244  
 ロジスティック回帰 (logistic regression) 242  
 ロジスティック変換 (logistic transformation) 233

<p><b>【A~E】</b></p> <p>AIC 116, 209</p> <p>AUC 112</p> <p>bag-of-words 237</p> <p>BIC 91, 117</p> <p>diag 33</p> <p>EM 67</p> <p><b>【F】</b></p> <p>F 値 110</p> <p>F 分布 (F-distribution) 21</p> <p>FDC (fault detection and classification) 8</p> <p><b>【G~I】</b></p> <p>GS 直交化 (GS orthogonalization) 53</p> <p>HOG 237</p> <p>IBIS 241</p> <p>IF-THEN ルール (IF-THEN rule) 3</p> <p><b>【K】</b></p> <p><math>k</math> 近傍法 (<math>k</math>-nearest neighbor method, method of <math>k</math>-nearest neighbors) 74, 121</p> <p><math>k</math> 平均法 (<math>k</math>-means) 84, 226</p>	<p><b>【L】</b></p> <p>LOF (local outlier factor) 76</p> <p>LU 分解 (LU factorization, LU decomposition) 249</p> <p><math>\ell_2</math> 正則化 (<math>\ell_2</math> regularization) 167</p> <p><math>L_2</math> 正則化 (<math>L_2</math> regularization) 167</p> <p><b>【M】</b></p> <p>MAP 167, 187</p> <p>Matlab 227</p> <p>MT システム (MT system) 49</p> <p>MT 法 (MT method) 49, 120</p> <p>MTGS 法 (MTGS method) 53</p> <p>MTS 49</p> <p><b>【N】</b></p> <p>NIPALS (nonlinear iterative partial least squares) 178</p> <p><b>【P】</b></p> <p>PLS 173</p>	<p><b>【Q】</b></p> <p>QR 分解 (QR factorization, QR decomposition) 53, 146</p> <p><b>【R】</b></p> <p>R 242</p> <p>RBF 展開 (RBF expansion) 79</p> <p>ROC 曲線 (receiver operating characteristic curve) 112</p> <p>RStudio 14</p> <p><b>【S】</b></p> <p>SIFT 237</p> <p>SN 比 (SN ratio) 50</p> <p>SPC 173</p> <p>subject to 99, 263</p> <p>SVM 243</p> <p><b>【T】</b></p> <p><math>t</math> 分布 (<math>t</math> distribution) 56</p> <p>TF-IDF 237</p> <p>Tr 38, 252</p> <p><b>【U】</b></p> <p>UL 分解 (UL factorization, UL decomposition) 249</p>
---	--	---

## Rコマンド索引

<p><b>【A】</b></p> <p>ada 242</p> <p>AIC 211</p> <p>ar 211</p> <p>as.matrix 51, 92, 171</p>	<p>as.numeric 172, 211</p> <p><b>【B】</b></p> <p>barplot 52, 110</p> <p>bkde2D 82</p>	<p><b>【C】</b></p> <p>car 11, 17, 24, 82, 91, 242</p> <p>Cars93 137, 150, 243</p> <p>cbind 42, 110, 171</p> <p>cc 185, 242</p>
--	--	---

CCA	185, 242	klaR	243	print	91, 138
cdens	92	kmeans	243		
chol	55	knn	242	<b>[Q]</b>	
class	242	knnx.dist	198, 242	qchisq	24, 25, 42
colMeans	43, 51	kpca	151	qr	53
colSums	83, 138, 139	ks	243	<b>[R]</b>	
contour	82, 101	ksvm	101	randomForest	243
<b>[D]</b>		<b>[L]</b>		rbind	101
datasets	242	lars	242	read.table	203
Davis	7, 17, 42, 59, 62, 82, 91, 93, 229, 242	length	63, 111, 171, 203, 211	rep	71, 101, 111, 203
density	82, 235	library	13, 242	river	195
dgamma	63	lines	25, 43, 51, 64, 172	rnorm	71, 101, 204
diag	139, 172	list	82	rotated	151
dnorm	71	lm.ridge	170, 243	rownames	110
dpik	82	<b>[M]</b>		rpart	243
dt	56	mahalanobis	243	runif	233
dtw	197, 242	MARSS	227, 243	<b>[S]</b>	
<b>[E]</b>		MASS	60, 137, 170, 243	sample	71, 240
ECG	211	Matrix	243	scale	101, 137, 172
eigen	137, 139	matrix	43, 51, 101, 243	sd	63
embed	198	mclust	11, 91, 243	seq	171, 204
evd	138	mean	24, 63	set.seed	204
<b>[F]</b>		mlbech	243	solve	43, 51, 172, 236
fitdistr	60, 243	mvpart	243	SOM	242
FNN	198, 242	mvtnorm	37, 243	sort	172
for	71, 111, 203	<b>[N]</b>		sqrt	71
<b>[G~I]</b>		NaiveBayes	243	stats	82, 211, 243
glm	243	notttem	211, 242	sum	71, 111, 172
glmnet	242	nrow	43, 51, 101	sunspots	242
hist	24, 233	nutrimouse	185, 242	svd	203
image	82, 101	<b>[O]</b>		<b>[T~Y]</b>	
is.element	137	order	64, 111, 138, 139	text	151
<b>[K]</b>		<b>[P]</b>		UScrime	170, 243
kde2d	243	plot	25, 42, 43, 51, 64, 83, 91, 101, 112, 137, 151, 172, 198, 235	var	24
kernlab	100, 150, 243			which	71
KernSmooth	82, 243			which.min	171
				y~.	171

— 著者略歴 —

1990年 国立苫小牧工業高等専門学校機械工学科卒業  
1993年 東北大学工学部機械工学科卒業  
2000年 東京大学大学院博士課程修了（物理学専攻）  
博士（理学）  
2000年 IBM 東京基礎研究所研究員  
2013年 IBM T. J. Watson Research Center,  
Senior Technical Staff Member  
現在に至る

入門 機械学習による異常検知 — Rによる実践ガイド —

Introduction to Anomaly Detection using Machine Learning

— A Practical Guide with R —

©Tsuyoshi Idé 2015

2015年3月13日 初版第1刷発行



検印省略

著者 井手 剛<sup>つよし</sup>  
発行者 株式会社 コロナ社  
代表者 牛来真也  
印刷所 三美印刷株式会社

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社

CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844・電話 (03)3941-3131(代)

ホームページ <http://www.coronasha.co.jp>

ISBN 978-4-339-02491-3 (金) (製本: 愛千製本所)

Printed in Japan



本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられております。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めておりません。

落丁・乱丁本はお取替えいたします