



パターン認識

早稲田大学講義 – 平成18年度

(独)産業技術総合研究所 脳神経情報研究部門
栗田多喜夫、赤穂昭太郎

講義の概要

最近、犯罪・テロの防止、交通事故の削減等の「安心・安全で快適な社会」を構築するための技術開発の重要性が指摘されている。これらの応用では「人間なら簡単にできるが今のコンピュータでは難しい」機能を実現することが重要である。

人間は、学習を通して、現実世界の多様で膨大な情報を類型的なパターンとして概念に対応付け、それらの関係を知識として蓄積することで、さまざまな状況に柔軟に対応できる。パターン認識は、人間が生存するための最も基本的な能力であり、知能の根幹をなしており、こうした課題の解決のための鍵を握っている。

また、インターネットや携帯電話等が急速に普及し、生活の様々な場面で情報技術が利用されるようになったが、パターン認識は、そうした情報機器と人間との自然なインタフェースを実現するための基本的な機能を提供する。さらには、インターネット上に分散的に蓄えられた大量のデータの中から意味のある情報を取り出すためのデータマイニングや遺伝子配列とその機能との関連性を抽出するバイオインフォマティクス等でも、パターン認識が重要な役割を担っている。

パターン認識の実現には、現実世界の曖昧さや不確かさを扱う必要がある。

本講義では、**パターン認識**および**機械学習**の話題について、確率統計的な視点から解説し、人間のような柔軟な知的情報処理システムを実現するための基礎の習得を目指す。

参考書・資料

- 参考書
 - R.O.Duda, P.E.Hart, and D.G.Stork, (尾上守夫監訳)、「パターン識別」、新技術コミュニケーションズ
 - 大津展之、栗田多喜夫、関田巖、「パターン認識—理論と応用—」、朝倉書店
 - C.M.Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
 - S.Theodoridis, K.Koutroumbas, Pattern Recognition, Academic Press, 1999.
 - T.Hastie, R.Tibshirani, and S.J.Friedman, The Elements of Statistical Learning – Data Mining, Inference, and Prediction --
- 参考資料
 - 「パターン認識とニューラルネットワーク」
 - 「サポートベクターマシン入門」
栗田のホームページ
<http://staff.aist.go.jp/takio-kurita/index-j.html>
からダウンロード可能

質問等

- 電子メール
takio-kurita@aist.go.jp
- 連絡先
〒305-8568
茨城県つくば市梅園1-1-1 つくば中央第2
産業技術総合研究所 脳神経情報研究部門
栗田 多喜夫
- 電話・FAX
電話 029-861-5838 FAX 029-861-5842

パターン認識とは

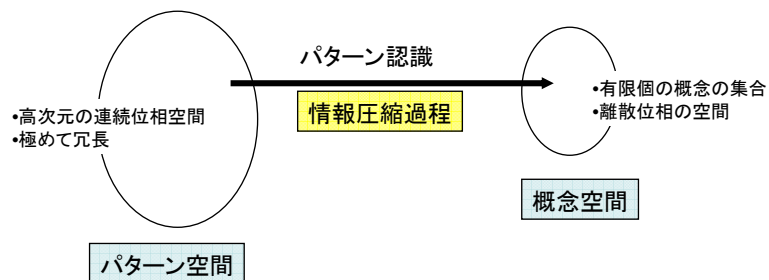
パターン認識の歴史

- パターン認識と人工知能
 - 認識や知能などの人間(生体)の脳の情報処理機能(知的情報処理機能)を解明し、それを機械(コンピュータ)で実現する試み
 - 情報処理技術に新たな概念を提供してきた
- 歴史
 - コンピュータ出現の初期
 - コンピュータは“万能機械”として、人間のあらゆる知的活動を代行してくれると期待(チェスなどのゲーム、作曲、自動翻訳、定理証明などへの応用)
 - ニューロンモデル(McCulloch & Pitts, 1943)、パーセプトロン(Rosenblatt, 1957)
 - 1960年代～
 - コンピュータへの入力装置として、文字・図形・音声などの機械による認識(パターン認識)の試み => まだまだ人間の能力には及ばない。
 - 1970年代～
 - 人工知能研究、第5世代コンピュータ(1982年～1992年)
 - 1980年代後半～
 - 誤差逆伝播学習法(Rumelhart, Hinton & Williams, 1986)、第2次ニューロブーム
 - リアルワールドコンピューティング(1992年～2002年)

パターン認識とは

- パターン認識

- 認識対象がいくつかの概念に分類出来るとき、観測されたパターンをそれらの概念(クラスあるいは類)のうちの一つに対応させる処理
 - 数字の認識: 入力パターンを10種類の数字のいずれかに対応させる
 - 顔画像の識別: 顔画像から誰であるかを推定する

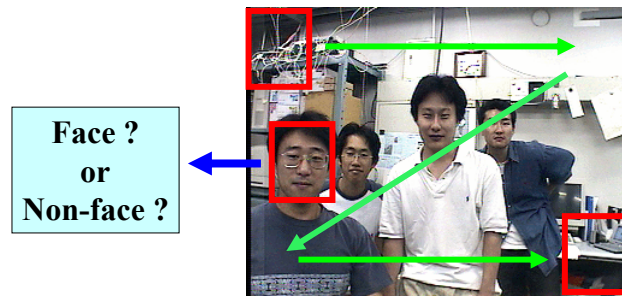


パターン認識問題の例: スпамフィルタ

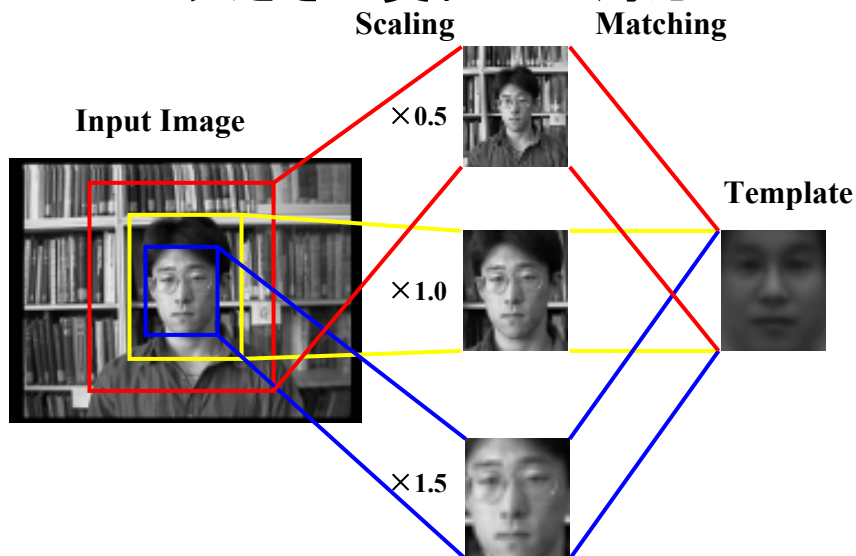
- スパムメールを検出して、自動削除する

- 特徴抽出
 - メール本文やヘッダにどのような単語が現れているかの頻度を計測し、それらをまとめて特徴ベクトルとする
- 訓練用のサンプルの作成
 - 過去のメールのデータベースから特徴ベクトルを計測し、そのメールがスパムかどうかを記録し、そのペアを訓練用サンプルデータとする
- 識別器の学習
 - 訓練用のサンプルを用いて識別器のパラメータを学習する
- 運用
 - 新たなメールから特徴ベクトルを計測し、それを識別器に入力し、その結果がスパムであれば、そのメールをスパムフォルダに移動する

画像中の顔の検出



大きさの変化への対応



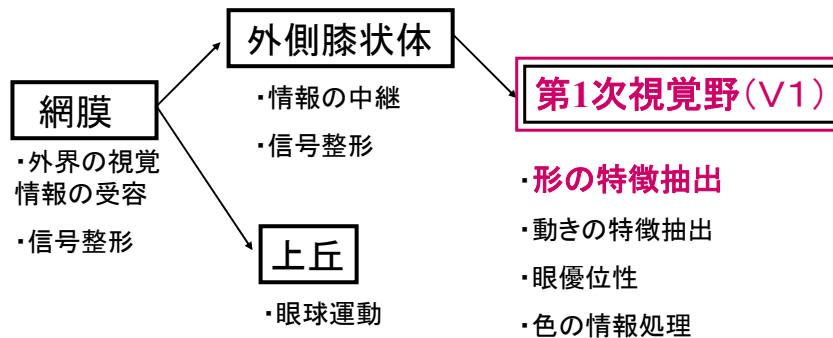
パターン認識問題の例

- ロボット
 - 顔、声から誰かを識別、音声から何を喋っているかを認識、手で触って、状態(柔らかい、硬い)を判定
- 車
 - 対向車や人の検出、運転者の状態(眠い、テンションがあがっている、...)
- 医療
 - 検査結果から病気を推定(肺がん)
- 軍事
 - ソナーデータから潜水艦かどうかを識別
- ワイン
 - 成分からワインの種類を識別

脳科学の進展

- 脳の視覚情報処理に関する知見
 - 網膜レベルからすでに情報が分化
 - 空間的な位置関係や動きに関する知覚---大脳皮質の視覚野から上に向かい頭頂連合野に至る経路
 - 視野内の物体が何かのパターン認識---視覚野から下の側頭連合野に至る経路
 - 視覚情報処理のための多くの専門分化された領野が存在
 - コラム構造
 - 眼優位性コラム(第1次視覚野V1)---左右どちらの芽からの情報を受け取るかでコラム構造を形成
 - 方位選択性コラム(第1次視覚野V1) --- 線分の方向に選択的に反応する細胞がコラム構造を形成
 - 三次元物体回転(TE野)---似た図形特徴に反応する細胞が三次元物体回転に対する見えの変化と整合性を持つような順序でコラム構造を形成
 - 運動方向性コラム(MT野)---視野内の刺激の方向に選択的に反応する細胞がコラム構造を形成

初期視覚情報処理

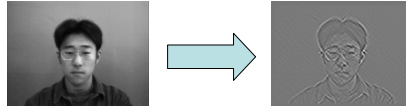
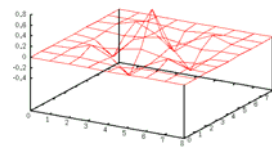
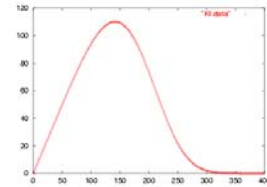


網膜および第一次視覚野の情報処理

- 何をやっているのか？
 - 入ってきた視覚情報からなるべく多くの情報を取り込み、しかも取り込んだ情報に含まれる冗長性をなるべく取り除くような情報処理を実現
- 網膜
 - 眼底に貼り付いている透明な神経組織
 - 外界の視覚情報を受け取り、局所的な情報処理の結果を神経パルス列に符号化して、視覚中枢に送り込む
 - 視細胞(photoreceptor)、水平細胞(horizontal cell)、アマクリン細胞(amacrine cell)、神経節細胞(ganglion cell)が整然と並んだ層構造
- 視覚第一次野(V1)
 - 6層構造をした後頭部にある大脳皮質の一部で、外側膝状体(LGN)からの入力は4C層に入る
 - 各ニューロンは受容野により規定される方向を持った直線状のコントラストに対して強い反応を示す(単純型細胞)
 - 光刺激の位置が方位に垂直方向に多少ずれても反応の強さが変化しないニューロンも存在する(複雑型細胞)

網膜での情報処理

- 自然画の局所的な自己相関
 - 風景や顔などの人工物を含まない自然画像の局所的な自己相関のパワースペクトルは空間周波数の2乗に反比例する(Field 1987)
- 神経節細胞の出力のパワースペクトル
 - 低周波では、平坦(コンスタント)(Atick等 1992)
 - 自己相関を空間的に無相関化していることに対応
 - 入力情報から空間的な冗長性を取り除く処理(whitening)
 - 高周波では、高周波成分を抑制
 - Whiteningによりノイズが増幅されることを防ぐ働き
- コントラストフィルタ



早稲田大学電気・情報生命工学科講義

独立行政法人 産業技術総合研究所

第一次視覚野での情報処理

- 情報抽出とスパース符号化
 - いくつかの基底ベクトルの線形結合により入力をなるべく近似し、しかも、その結合係数がなるべくスパースになるような基準で基底ベクトルを求めると、第一次視覚野の単純型細胞の特性と似た特徴が得られる(Olshausen & Field, 1996)
- 独立成分の抽出
 - 独立成分分析(ICA)を用いて、Olshausen & Fieldの結果と同様な結果が得られる(Bell & Sejnowski, 1997)

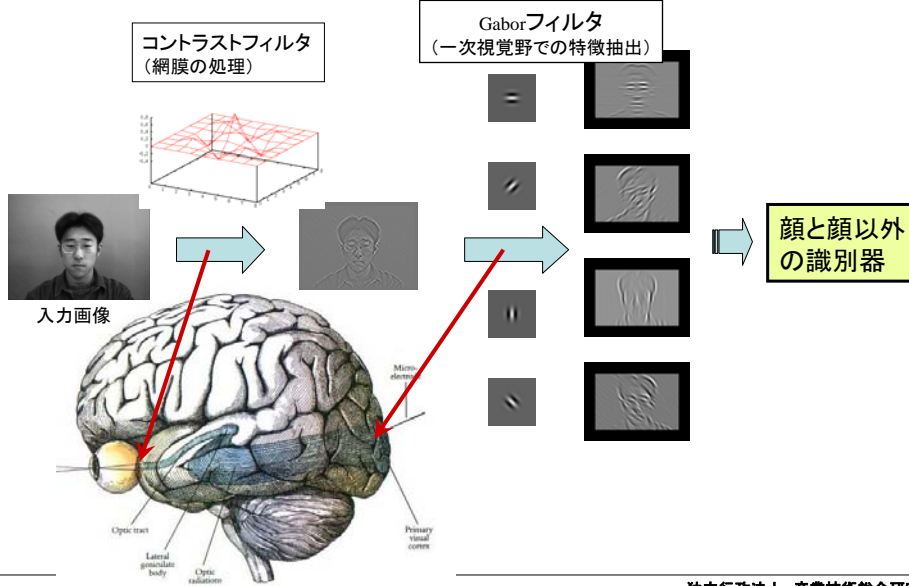
入ってきた視覚情報からなるべく多くの情報を取り込み、しかも取り込んだ情報に含まれる冗長性をなるべく取り除くような情報処理を実現

- これは、入力情報と取り込む機能の実現するための最も自然な動作原理

早稲田大学電気・情報生命工学科講義

独立行政法人 産業技術総合研究所

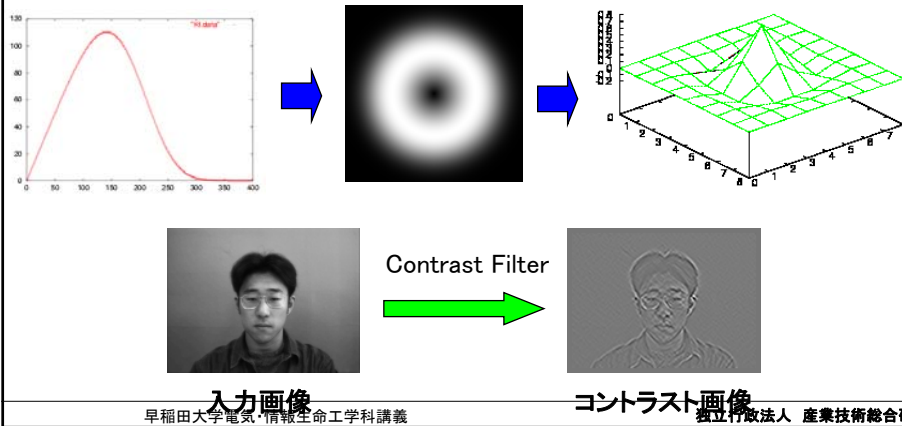
顔検出のためのセンシングシステムの例



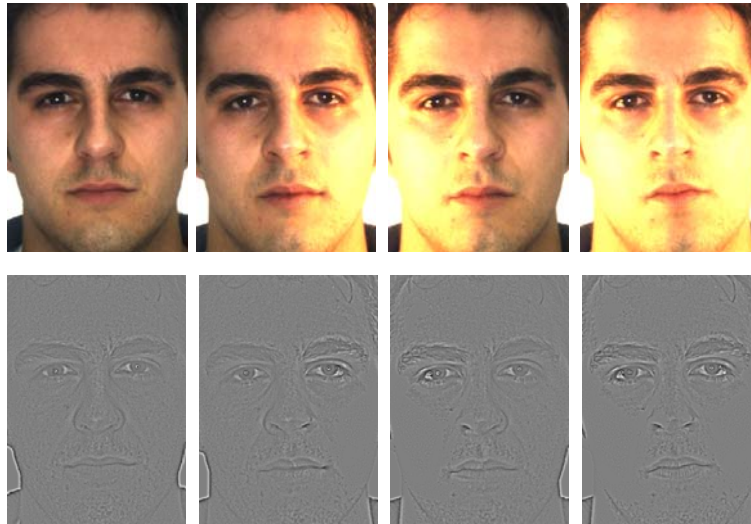
コントラストフィルタ

- 網膜のガングリオン細胞の受容野に類似 [Atick92, Olshausen97]

$$K(f) = W(f)L(f) = f \exp \left\{ - \left(\frac{f}{f_0} \right)^4 \right\}$$

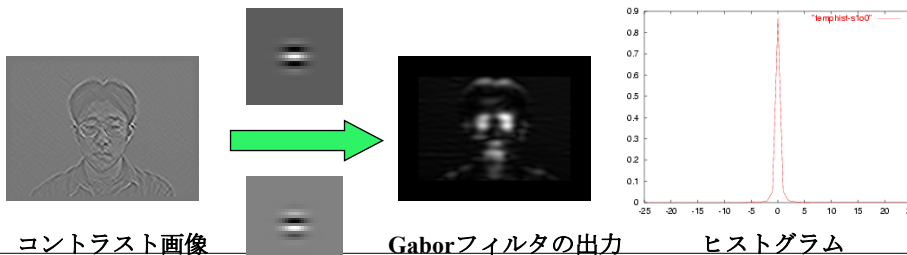


コントラストフィルタの 明るさの変化に対する頑健性



Gaborフィルタ

- V1野の単純型細胞の受容野特性に類似 [Jones87]
- 顔(対象)認識への有効性が報告されている [Malsburg93]
 - Sparse coding : 鋭い選択性を持つ細胞集団の発火により情報を表現 [Olshausen96]
 - 自然画像のICA [Bell96] → Gabor-likeフィルタ
 - ➡ 各方位のGaborフィルタ: 確率的独立性が高い
- 実験: 8方向のGaborフィルタ(9x9画素)を利用



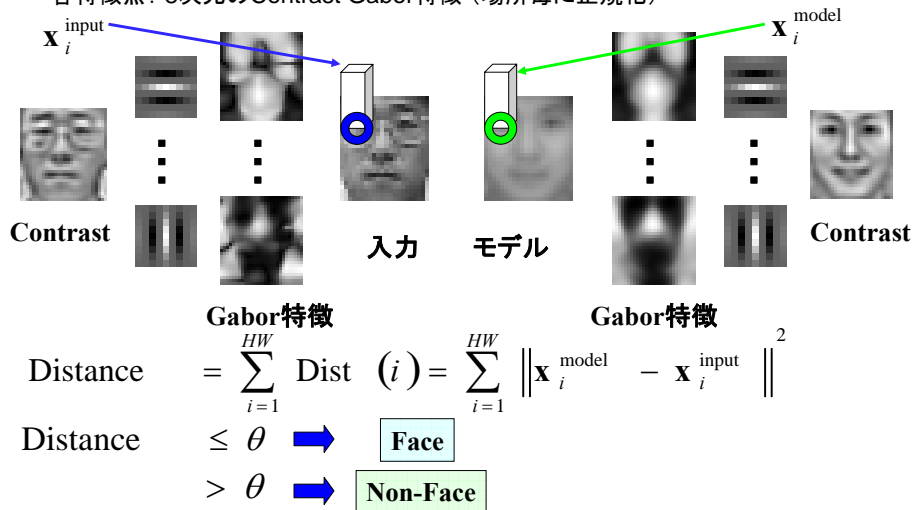
コントラスト画像

Gaborフィルタの出力

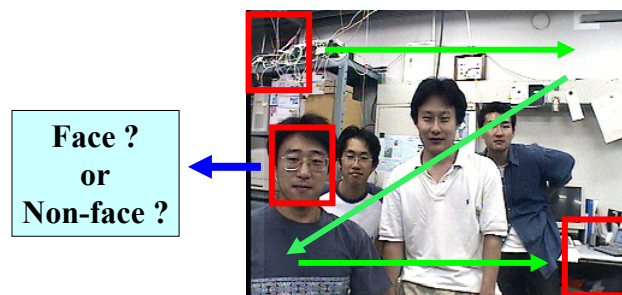
ヒストグラム

顔と顔以外のものとの識別

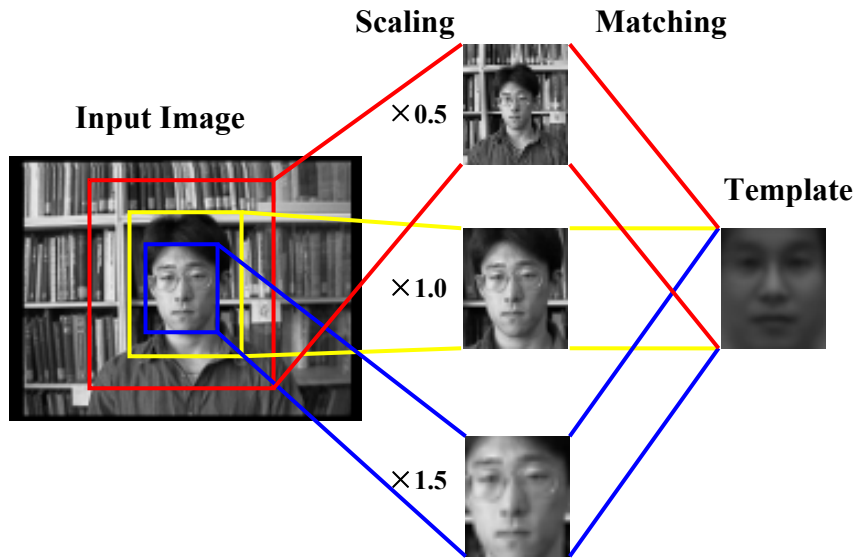
- 識別器 : モデルとのマッチング
- 各特徴点 : 8次元のContrast Gabor特徴 (場所毎に正規化)



画像中の顔の検出



大きさの変化への対応



特定の人を顔として検出

テンプレート

約200枚の顔画像に対して正しく検出できた

平均顔 (検出された約200枚の顔画像から作成した)

平均顔 (相関マッチングで検出した顔画像で作成。顔の検出率31.7%)

1996年に撮影

1997年に撮影

1998年に撮影

1999年に撮影 (暗い)

1999年に撮影 (隠れあり)

早稲田大学電気・情報生命工学科講義

独立行政法人 産業技術総合研究所

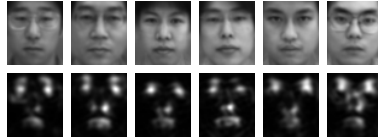
顔検出 + 個人識別

顔検出
(多数の人の平均顔)



162枚の顔画像に対して、
99.4%の検出+個人識別率

個人識別
(個人の平均顔)



1996年に撮影



かなり暗い



めがねをかけた



数学的準備

線形代数と確率統計の復習

ベクトル、行列

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

転置

$$\mathbf{x}^T = (x_1 \quad x_2 \quad \cdots \quad x_d)$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

正方行列、対象行列、単位行列

- 正方行列
m=nの行列

- 対象行列

$$A = A^T$$

- 対角行列

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_m \end{pmatrix} = \text{diag}(d_1 \quad d_2 \quad \cdots \quad d_m)$$

- 単位行列

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = [\delta_{ij}]$$

行列とベクトルの積

$$\begin{aligned}
 \mathbf{y} &= A\mathbf{x} \\
 &= \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \\
 &= \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}
 \end{aligned}$$

ベクトルの内積、ノルム

内積

$$\mathbf{x}^T \mathbf{y} = (x_1 \quad x_2 \quad \cdots \quad x_m) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \sum_{i=1}^m x_i y_i = \mathbf{y}^T \mathbf{x}$$

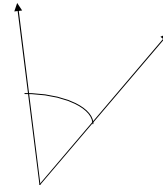
ユークリッドノルム(長さ)

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}} = \sqrt{\sum_{i=1}^m x_i^2}$$

ベクトルのなす角、直交、平行

ベクトルのなす角

$$\cos \theta = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$



直交

$$\mathbf{x}^T \mathbf{y} = 0$$

コーシー・シュワルツの不等式

$$\|\mathbf{x}^T \mathbf{y}\| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

平行

$$\|\mathbf{x}^T \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\|$$

線形独立、ベクトル空間

- 線形独立性
 - どのベクトルも他のベクトルの線形結合として表せない場合
- ベクトル空間
 - d個の線形独立なベクトルは、d次元のベクトル空間を張る
 - d次元ベクトル空間の任意のベクトルは、d個の線形独立なベクトルの線形結合で表せる。

ベクトルの積

$$\mathbf{xy}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} (y_1 \quad y_2 \quad \cdots \quad y_m) = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_m \\ \vdots & \vdots & \vdots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_m \end{pmatrix}$$

勾配

関数

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_m)$$

勾配(微分)

$$\nabla f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_m} \end{pmatrix}$$

ベクトル関数とその微分

ベクトル関数

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix} : \mathbb{R}^m \rightarrow \mathbb{R}^n$$

ベクトル関数の微分

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \frac{\partial f_1(\mathbf{x})}{\partial x_m} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \frac{\partial f_2(\mathbf{x})}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_2} & \frac{\partial f_n(\mathbf{x})}{\partial x_m} \end{pmatrix} : (n \times m)$$

行列のベクトルに関する微分

$$\frac{\partial (A\mathbf{x})}{\partial \mathbf{x}} = A$$

$$\frac{\partial (\mathbf{y}^T \mathbf{x})}{\partial \mathbf{x}} = \frac{\partial (\mathbf{x}^T \mathbf{y})}{\partial \mathbf{x}} = \mathbf{y}$$

$$\frac{\partial (\mathbf{x}^T A \mathbf{x})}{\partial \mathbf{x}} = (A + A^T) \mathbf{x}$$

Aが対称行列の場合

$$\frac{\partial (\mathbf{x}^T A \mathbf{x})}{\partial \mathbf{x}} = 2A\mathbf{x}$$

逆行列、固有値、固有ベクトル

- 正則な(行列式が0でない)正方行列には、逆行列が存在

$$AA^{-1} = A^{-1}A = I$$

- 正方行列の固有値、固有ベクトル

$$A\mathbf{x} = \lambda\mathbf{x}$$

最適化問題の解法

関数

$$f(\mathbf{x})$$

の極値(最小値、最大値)を求める

- 最急降下法
 - 初期値からはじめて、微分方向に逐次更新

$$\mathbf{x}^{t+1} \leftarrow \mathbf{x}^t - \alpha \frac{\partial(f(\mathbf{x}))}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}^t}$$

制約条件のある最適化

問題: 制約条件

$$g(\mathbf{x}) = 0$$

のもとで、関数

$$f(\mathbf{x})$$

の極値(最小値、最大値)を求める

ラグランジュ関数

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

を考えることで、制約の無い問題に帰着させて解く方法

λ : ラグランジュの未定乗数

離散確率

- 有限個の離散的な値をとる変数があるとき、ある値を取る確率(例、サイコロ)

$$\mathcal{X} = \{v_1, v_2, \dots, v_m\}$$

$$p_i = \Pr[x = v_i], \quad (i = 1, \dots, m)$$

確率の条件

$$p_i \geq 0, \quad \sum_{i=1}^m p_i = 1$$

平均値、期待値

平均値

$$E[x] = \mu = \sum_{x \in \mathbb{N}} xP(x) = \sum_{i=1}^m v_i p_i$$

期待値

$$E[f(x)] = \sum_{x \in \mathbb{N}} f(x)P(x)$$

期待値の線形性

$$E[a_1 f_1(x) + a_2 f_2(x)] = a_1 E[f_1(x)] + a_2 E[f_2(x)]$$

2次のモーメント、分散

2次のモーメント

$$E[x^2] = \sum_{x \in \mathbb{N}} x^2 P(x)$$

分散

$$\begin{aligned} \text{Var}[x] &= \sigma^2 = E[(x - \mu)^2] = \sum_{x \in \mathbb{N}} (x - \mu)^2 P(x) \\ &= E[x^2] - (E[x])^2 \end{aligned}$$

2変数の場合

$$\mathcal{N}_1 = \{v_1, v_2, \dots, v_m\}, \mathcal{N}_2 = \{\omega_1, \omega_2, \dots, \omega_n\}$$

$$p_{ij} = \Pr[x = v_i, y = \omega_j] = P(x, y)$$

例: 2つのサイコロ

確率の条件

$$p_{ij} \geq 0, \quad \sum_{i=1}^m \sum_{j=1}^n p_{ij} = 1$$

周辺分布

$$P_y(y) = \sum_{x \in \mathcal{N}_1} P(x, y)$$

$$P_x(x) = \sum_{y \in \mathcal{N}_2} P(x, y)$$

統計的独立性

- 以下の式が成り立つとき、ふたつの変数は統計的に独立という

$$P(x, y) = P_x(x)P_y(y)$$

2つのサイコロの例

	1	2	3	4	5	6	
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	1/6	1/6	1/6	1/6	1/6	1/6	

2変数関数の期待値

期待値

$$E[f(x, y)] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} f(x, y) P(x, y)$$

平均値

$$\mu_x = E[x] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} x P(x, y)$$

$$\mu_y = E[y] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} y P(x, y)$$

2変数関数の分散

分散

$$\sigma_x^2 = E[(x - \mu_x)^2] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} (x - \mu_x)^2 P(x, y)$$

$$\sigma_y^2 = E[(y - \mu_y)^2] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} (y - \mu_y)^2 P(x, y)$$

共分散

$$\sigma_{xy} = E[(x - \mu_x)(y - \mu_y)] = \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} (x - \mu_x)(y - \mu_y) P(x, y)$$

相関係数

相関係数

共分散を標準偏差で正規化したもので-1から1の間の値を取る

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

相関係数が1なら、正の相関が最大
相関係数が-1なら、負の相関が最大

相関係数が0なら、相関が無い(無相関)

統計的独立なら、無相関

条件付確率

条件付確率

$$P(x|y) = \frac{P(x, y)}{P(y)} = \frac{P(x)P(y|x)}{P(y)}$$

サイコロの例([奇数、偶数]、[1の目が出る、1以外の目が出る])

	奇数	偶数	P(y)
1が出る	1/6	0	1/6
1以外が出る	1/3	1/2	5/6
P(x)	1/2	1/2	

平均ベクトル、分散共分散行列

平均ベクトル

$$E[\mathbf{x}] = \boldsymbol{\mu} = \sum_{\mathbf{x} \in \mathcal{N}^m} \mathbf{x}P(\mathbf{x})$$

分散共分散行列

$$\Sigma = E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] = E[\mathbf{x}\mathbf{x}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T$$

連続な確率変数の場合

確率密度分布 $p(x)$

$$p(x) \geq 0 \quad \int_{-\infty}^{\infty} p(x)dx = 1$$

期待値、平均、分散

$$E[f(x)] = \int_{-\infty}^{\infty} f(x)p(x)dx$$

$$\mu = E[x] = \int_{-\infty}^{\infty} xp(x)dx$$

$$\text{Var}[x] = \sigma^2 = E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx$$

多変数の場合

確率密度分布 $p(\mathbf{x})$

$$p(\mathbf{x}) \geq 0 \quad \int_{-\infty}^{\infty} p(\mathbf{x}) d\mathbf{x} = 1$$

期待値、平均、分散

$$E[f(\mathbf{x})] = \int_{-\infty}^{\infty} f(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}$$

$$\boldsymbol{\mu} = E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x}) d\mathbf{x}$$

$$\boldsymbol{\Sigma} = E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] = \int_{-\infty}^{\infty} (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T p(\mathbf{x}) d\mathbf{x}$$

統計的独立性、条件付確率

統計的独立性

$$p(\mathbf{x}) = \prod_{i=1}^m p(x_i)$$

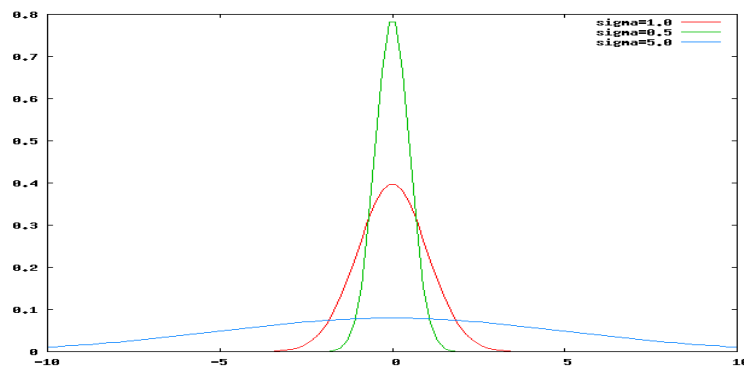
条件付確率

$$p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(x)p(y|x)}{\int_{-\infty}^{\infty} p(x)p(y|x)dx}$$

正規分布

正規分布

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$



正規分布の平均、分散

平均、分散

$$E[1] = \int_{-\infty}^{\infty} p(x) dx = 1$$

$$E[x] = \int_{-\infty}^{\infty} xp(x) dx = \mu$$

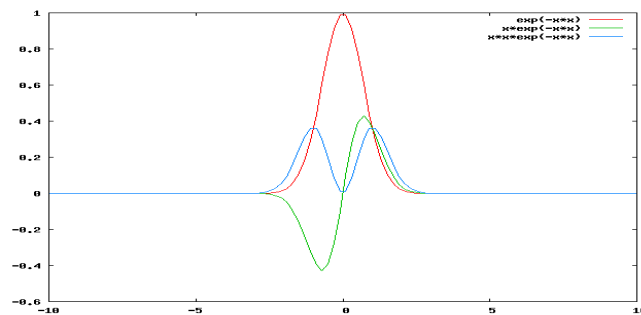
$$E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x) dx = \sigma^2$$

積分の公式

$$\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}$$

$$\int_{-\infty}^{\infty} x \exp(-x^2) dx = 0$$

$$\int_{-\infty}^{\infty} x^2 \exp(-x^2) dx = \frac{\sqrt{\pi}}{2}$$



マハラノビス距離

マハラノビス(Mahalanobis)距離
標準偏差で正規化した平均までの距離

$$d = \frac{|x - \mu|}{\sigma}$$

正規分布の標準形
平均0、分散1の正規分布

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}$$

多変量正規分布

正規分布

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}$$

平均、分散

$$E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x}p(\mathbf{x})d\mathbf{x} = \boldsymbol{\mu}$$

$$E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] = \int_{-\infty}^{\infty} (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T p(\mathbf{x})d\mathbf{x} = \Sigma$$

マハラノビス距離(多変量の場合)

マハラノビス(Makaranobis)距離

$$d^2 = (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})$$

1 限目終了