パターン認識

早稲田大学講義 – 平成20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、質の高い生活の支援等の「安心・安全で快適な社会」を実現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。そのため確率統計的な手法が重要となる。

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。

講義の内容

- **前半 (栗田担当)**
  - http://staff.aist.go.jp/takio-kurita/index-j.html
  - **パターン認識とは、数学的準備**
  - **統計的決定理論 (ベイズ決定理論、確率密度関数の推定)**
  - **線形識別関数の学習 (パーセプトロン、ロジスティック回帰、多層パーセプトロン)**
  - **汎化性 (汎化性能の評価、特徴選択)**
  - **統計的特徴抽出 (回帰分析、主成分分析、判別分析)**
  - **クラスタリング**

- **後半 (赤穂担当)**
  - http://www.neurosci.aist.go.jp/~akaho/waseda/

AIST 脳神経情報研究部門

参考書·資料

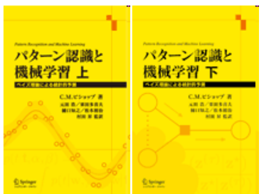
- **参考書**
  - C.M.Bishop, Pattern Recognition and Machine Learning,Springer, 2006.
    (訳本:パターン認識と機械学習 (上)、(下)、シュプリンガー·ジャパン)
  - R.O.Duda, P.E.Hart, and D.G.Stork, (尾上守夫監訳)、「パターン識別」、新技術コミュニケーションズ
  - 大津展之、栗田多喜夫、関田巌、「パターン認識―理論と応用―」、朝倉書店
  - S.Theodoridis, K.Koutroumbas, Pattern Recognition, Academic Press, 1999.
  - T.Hastie, R.Tibshirani, and SJ.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



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

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

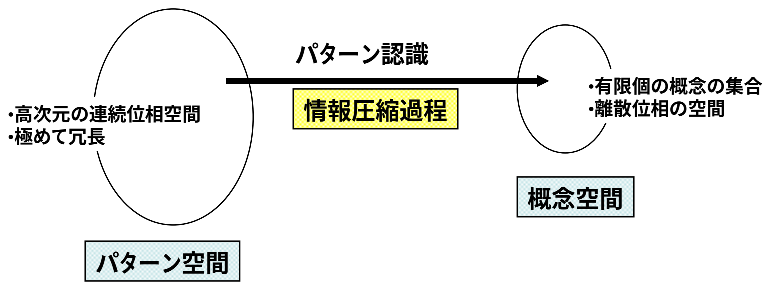
パターン認識とは

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）のうちのひと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：スパムフィルタ

- スパムメイルを検出して、自動削除する
  - 特徴抽出
    - メイ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
    - 過去のメイルのデータベースから特徴ベクトルを計測し、そのメイ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
    - 新たなメイ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、そのメイルをスパムフォールダに移動する

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

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

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

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

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

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**

$$A = A^T$$

- **単位行列**

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

- **対角行列**

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

行列とベクトルの積

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

$$= \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix}$$

ベクトルの内積、ノルム

内積

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

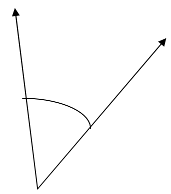
ユークリッドノルム (長さ)

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

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

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

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

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix} = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_m \\ \vdots & \vdots & \vdots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_m \end{pmatrix}$$

勾配

関数

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

勾配（微分）

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

ベクトル関数とその微分

ベクトル関数

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

ベクトル関数の微分

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

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

$$\mathbf{x}^{t+1} \Leftarrow \mathbf{x}^{t} - \alpha \left. \frac{\partial (f(\mathbf{x}))}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}^{t}}$$

最適化問題の解法

- ニュートン法
  - Talor展開

$$f(\mathbf{x}^{t} + \boldsymbol{\delta}) \approx f(\mathbf{x}^{t}) + \left(\frac{\partial f}{\partial \mathbf{x}}\right)\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\delta}^{T}\left(\frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^{T}}\right)\boldsymbol{\delta}$$

  - これを最小とするには、変移で微分して0とおくと、

$$\left(\frac{\partial f}{\partial \mathbf{x}}\right) + \left(\frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^{T}}\right)\boldsymbol{\delta} = 0 \qquad \boldsymbol{\delta} = -\left(\frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^{T}}\right)^{-1}\left(\frac{\partial f}{\partial \mathbf{x}}\right)$$

  - 更新式は

$$\mathbf{x}^{t+1} \Leftarrow \mathbf{x}^{t} - \left(\frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^{T}}\right)^{-1} \left. \frac{\partial (f(\mathbf{x}))}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}^{t}}$$

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）**

$$\aleph = \{v_1, v_2, \ldots, v_m\}$$
$$p_i = \Pr[x = v_i], \quad (i = 1, \ldots, m)$$

確率の条件

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

平均値、期待値

平均値

$$\mathrm{E}[x] = \mu = \sum_{x \in \aleph} x P(x) = \sum_{i=1}^{m} v_i p_i$$

期待値

$$\mathrm{E}[f(x)] = \sum_{x \in \aleph} f(x) P(x)$$

期待値の線形性

$$\mathrm{E}[a_1 f_1(x) + a_2 f_2(x)] = a_1 \mathrm{E}[f_1(x)] + a_2 \mathrm{E}[f_2(x)]$$

2次のモーメント、分散

2次のモーメント

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

分散

$$\mathrm{Var}[x] = \sigma^2 = \mathrm{E}[(x-\mu)^2] = \sum_{x \in \aleph} (x-\mu)^2 P(x)$$

$$= \mathrm{E}[x^2] - (\mathrm{E}[x])^2$$

2変数の場合

$$\aleph_1 = \{v_1, v_2, \ldots, v_m\}, \aleph_2 = \{\omega_1, \omega_2, \ldots, \omega_n\}$$
$$p_{ij} = \Pr[x = v_i, y = \omega_j] = P(x, y)$$

例：2つのサイコロ

確率の条件

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

周辺分布

$$P_y(y) = \sum_{x \in \aleph_1} P(x, y)$$
$$P_x(x) = \sum_{y \in \aleph_2} P(x, y)$$

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

$$\mathrm{E}[f(x,y)] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_{21}} f(x,y)P(x,y)$$

平均値

$$\mu_x = \mathrm{E}[x] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_2} xP(x,y)$$

$$\mu_y = \mathrm{E}[y] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_2} yP(x,y)$$

2変数関数の分散

分散

$$\sigma_x{}^2 = \mathrm{E}[(x-\mu_x)^2] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_2}(x-\mu_x)^2 P(x,y)$$

$$\sigma_y{}^2 = \mathrm{E}[(y-\mu_y)^2] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_2}(y-\mu_y)^2 P(x,y)$$

共分散

$$\sigma_{xy}{}^2 = \mathrm{E}[(x-\mu_x)(y-\mu_y)] = \sum_{x\in\aleph_1}\sum_{y\in\aleph_2}(x-\mu_x)(y-\mu_y)P(x,y)$$

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

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

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

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

分散共分散行列

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

連続な確率変数の場合

確率密度分布 $p(x)$

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

期待値、平均、分散

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

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

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

多変数の場合

確率密度分布 $p(\mathbf{x})$

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

期待値、平均、分散

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

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

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

統計的独立性、条件付確率

統計的独立性

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

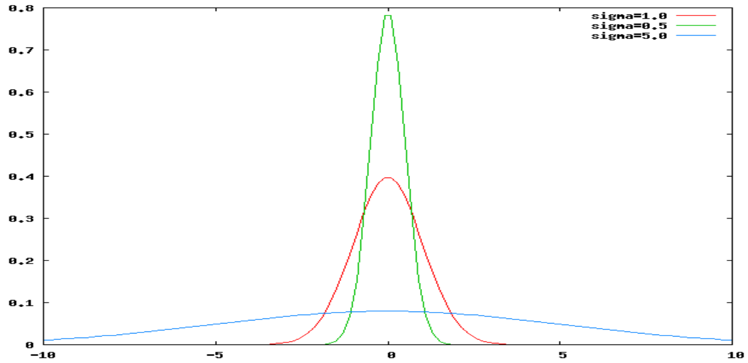
条件付確率

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

正規分布

正規分布

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



AIST 脳神経情報研究部門

正規分布の平均、分散

平均、分散

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

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

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

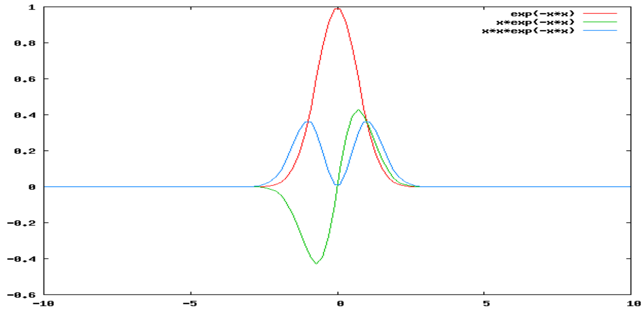
早稲田大学電気･情報生命工学科講義　独立行政法人 産業技術総合研究所

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(Makaranobis)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

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

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

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