ロジスティック回帰・分類モデルの評価

樋口さぶろお https://hig3.net

龍谷大学 先端理工学部 数理・情報科学課程

多変量解析☆演習　L06(2026-10-29 Thu)

最終更新: Time-stamp: "2026-11-04 Wed 23:11 JST hig"

今日の目標

- 回帰と分類の違いを説明できる
- 分類モデルを ROC と混同行列で評価できる
- ロジスティック回帰の確率モデルを説明できる

ここまで来たよ

6. ロジスティック回帰・分類モデルの評価
   - 単回帰・重回帰では変数選択重要
   - 分類問題
   - ロジスティック回帰の統計モデル
   - 分類問題の精度評価

重回帰は単回帰の上位互換か？

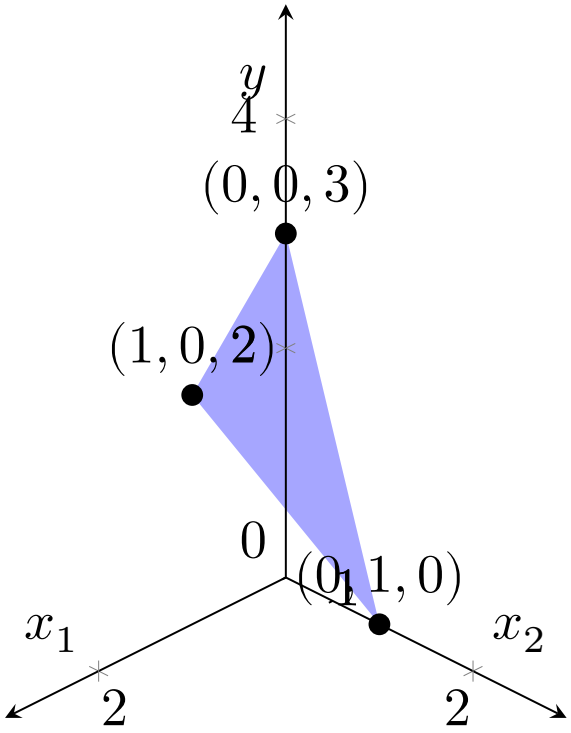
偏回帰係数は単回帰係数に等しいか？
実は，次は異なる

- x_1 で単回帰した単回帰係数 β_1，x_2 で単回帰した単回帰係数 β_2
- x_1, x_2 で重回帰した偏回帰係数 β_1, β_2

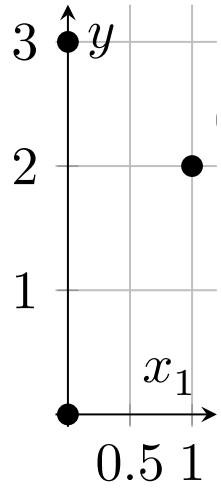
- 誤：「説明変数になるべくぜんぶ入れて重回帰を 1 回やっとけばいい」
- 誤：「単回帰繰り返せば偏回帰係数が全部わかる」
- 正：「個数が少ない割に決定係数が大きい（自由度調整済決定係数が大きい）ように説明変数を選ぶ」

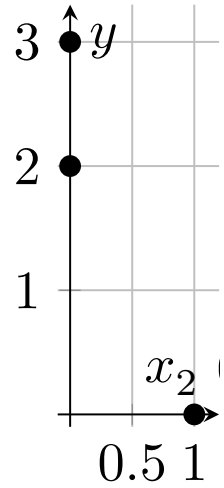
これは，x_1, x_2 の選び方に相関があるため，$X^\top X$ が対角行列でないため．
細いリボン状のデータ分布を想像すると納得いくが，ここでは，3 個からなる具体的データで考察

$$\begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} s + \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} t, \quad y = 3 - 1x_1 - 3x_2 \rightsquigarrow \beta_1' = -1, \beta_2' = -3$$



(x_1, y) 平面への射影　(x_2, y) 平面への射影





$y = 3 - x_1 - 3x_2$

$y = \beta_0 + \beta_1 x_1$
$\rightsquigarrow \beta_1 > 0$

$y = \beta_0 + \beta_2 x_2$
$\rightsquigarrow \beta_2 < 0$

平面の方程式のセンスだけでは済まない
データ点が (x_1, x_2) のどこに分布するか（リボン的，多重共線性的視点）
データ点が平面からどのくらい離れるか（最小二乗法的視点）

ここまで来たよ

6. ロジスティック回帰・分類モデルの評価
   - 単回帰・重回帰では変数選択重要
   - 分類問題
   - ロジスティック回帰の統計モデル
   - 分類問題の精度評価

予測問題の回帰タイプと分類タイプ

線形重回帰の扱うデータ
連続値を予測

	説明変数		目的変数
	x_1	x_2	y
1	10.3	1.1	21.1
2	12.1	3.1	22.9
⋮			
n	19.8	2.5	41.2

機械学習の言葉では
教師あり, 回帰=目的変数が連続値
y は小さい正規ノイズを含む.

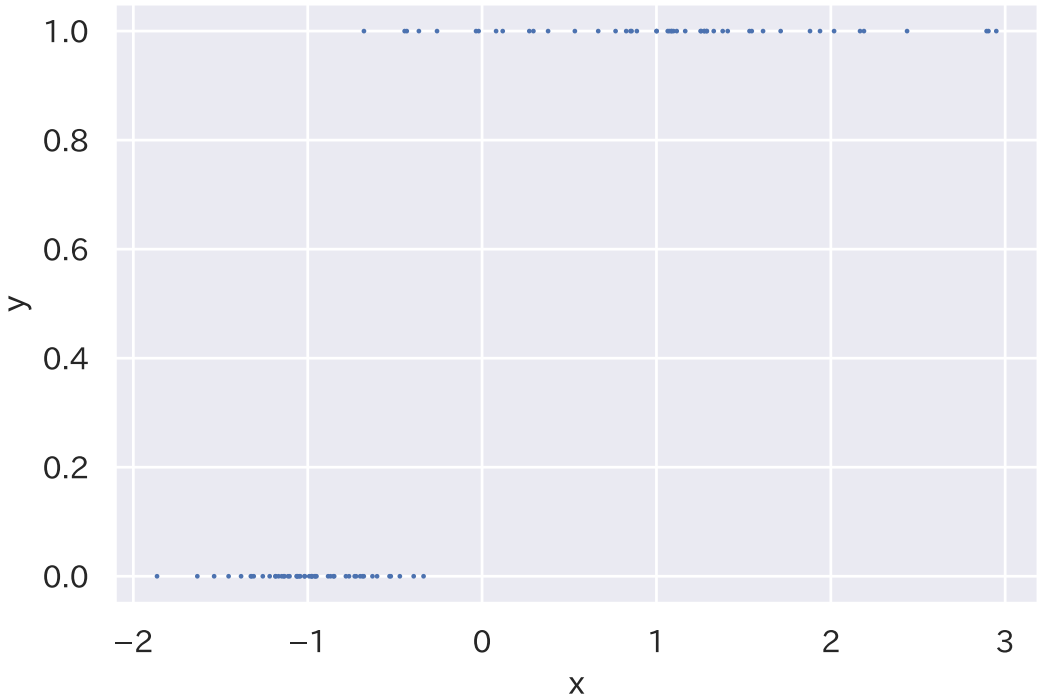
ロジスティック回帰で扱うデータ
離散値を予測=分類

	説明変数		目的変数	ダミー
	x_1	x_2	y	変数 y
1	10.3	1.1	A	0
2	12.1	3.1	B	1
⋮				
n	19.8	2.5	A	0

機械学習の言葉では
教師あり, 分類=目的変数がカテゴリ変数 ≃ 離散
y は確率的（ベルヌーイ試行）

ロジスティック回帰で対応できる学習データ

x 連続値
y 間隔に意味のない有限個 $\simeq$ 離散型, 今日は A→0, B→1 に限定
x の値だけで y が決まるわけではない．確率的な振る舞い．



多変量解析の手法の分類

教師あり

タイプ	教師 Y	手法・モデル
回帰	連続値	線形単回帰 L02 線形重回帰 L03,L04
	非負整数	ポアソン回帰
分類	2 値（ラベル）	ロジスティック回帰 L05
	多値（ラベル・カテゴリ）	ベイズ分類 2 次判別分析 (QDA)

時系列解析

L12 — とりあえず別世界

教師なし

タイプ	手法・モデル
クラスタリング	k-means 法
次元圧縮	主成分分析 (PCA) 因子分析 (Factor Analysis)

ここまで来たよ

6. ロジスティック回帰・分類モデルの評価
   - 単回帰・重回帰では変数選択重要
   - 分類問題
   - ロジスティック回帰の統計モデル
   - 分類問題の精度評価

ロジスティック回帰

定義 (ロジスティック単回帰モデル)

$$P(Y=1) = S(\boldsymbol{\beta}^\top \boldsymbol{x}), \quad \text{つまり} \quad Y \sim Bin(1, S(\boldsymbol{\beta}^\top \boldsymbol{x}))$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} 1 \\ x_1 \end{pmatrix}, \quad \boldsymbol{\beta}^\top \boldsymbol{x} = \beta_0 + \beta_1 \cdot x_1$$

後者は，線形重回帰モデルの，$Y \sim N(\boldsymbol{\beta}\boldsymbol{x}, \sigma^2)$ という表現に対応．
線形回帰: ロジスティック回帰=正規: 二項=母平均値: 母比率

定義 (ベルヌーイ分布 久保川 統計学入門 §4.3.2 | 確率統計 I(2026)L08)

$n = 1$ の二項分布
離散型確率変数 $X \sim Bin(1, p)$ とは，X が次の確率関数を持つこと．

$$p(x) = \begin{cases} p & (x = 1) \\ 1 - p & (x = 0) \end{cases}$$

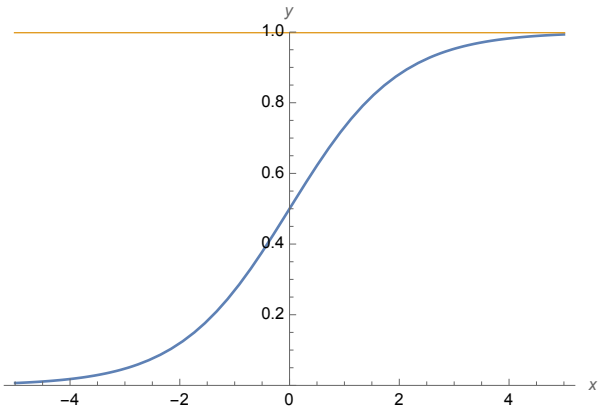
定義 (シグモイド関数)

$S : \mathbb{R} \to (0, 1)$

$$S(x) = \frac{1}{1 + \mathrm{e}^{-x}}$$

$\mathrm{expit}(x), \sigma(x)$ などとも書かれる.

x	$-\infty$		0		$+\infty$
y'	$+$	$+$	$\frac{1}{4}$	$+$	$+$
y	0	↗	$\frac{1}{2}$	↗	$+1$



L06-Q1

ロジスティック関数

$S(x) = \frac{1}{1+\mathrm{e}^{-x}}$ とする.

1. $S(x) = \frac{1}{2}$ となる x, $\frac{\mathrm{d}S}{\mathrm{d}x}$ が最大となる x を求めよう.
2. $f(x) = S(\beta_0 + \beta_1 x) = \frac{1}{2}$ となる x, $\frac{\mathrm{d}f}{\mathrm{d}x}$ が最大となる x を求めよう.

定義 (ロジスティック関数)

シグモイド関数を移動拡大縮小したもの $f(x) = S(\boldsymbol{\beta}^\top \boldsymbol{x})$ をロジスティック関数という.

$\beta_1 > 0$ とする.

x	$-\infty$		$-\beta_0/\beta_1$		$+\infty$
$\frac{\mathrm{d}f}{\mathrm{d}x}$	$+$	$+$	$\frac{\beta_1}{4}$	$+$	$+$
$f(x) = \frac{1}{1+\mathrm{e}^{-\boldsymbol{\beta}^\top \cdot \boldsymbol{x}}}$	0	↗	$\frac{1}{2}$	↗	$+1$

余談：一般化線形モデル

両辺の逆関数をとって，右辺を 1 次関数にした
$S^{-1}(y) = \beta_0 + \beta_1 x$ という書き方の方が趣味がよい．

誤差構造	リンク関数	説明変数の個数は p
$Y \sim N(\mu, \sigma^2)$	$\mathrm{id}(\mu) =$	$\boldsymbol{\beta}^\top \cdot \boldsymbol{x}$
$Y \sim Bin(1, p)$	$\mathrm{logit}(p) =$	$\boldsymbol{\beta}^\top \cdot \boldsymbol{x}$

- 恒等関数 identity function $\mathrm{id}(\mu) = \mu$
- ロジット関数 logit function $\mathrm{logit}(p) = \log \frac{p}{1-p}$ は $S(x)$ の逆関数．
  - $\frac{p}{1-p}$: オッズ比．0 と ∞ の間の値を取る
  - $\log \frac{p}{1-p}$: 対数オッズ比．$-\infty$ と $+\infty$ の間の値を取るから，これが x の 1 次式で決まるとしてもおかしくない

ロジスティック回帰はもっとも 1:1 次関数 対 ロジスティック関数

$$p(x) = P(Y = 1) = \beta_0 + \beta_1 \cdot x_1$$

じゃだめなの?

$\rightsquigarrow p = \beta_0 + \beta_1 x$ は,
$-\infty < x < +\infty$ のとき $-\infty < p < +\infty$.
確率のとる範囲 $0 \leq p \leq 1$ に収まらない…
$x \in (-\infty, \infty)$ を $y \in (0, 1)$ に写す写像ある?

ロジスティック回帰はもっとも 2: 母平均値 対 母比率

	説明変数なし	説明変数あり
母平均値	$Y \sim N(\mu, \sigma^2)$ の母平均値 μ の推定 $\overline{Y} = \frac{1}{n}(Y_1 + \cdots)$	線形回帰 $Y = \beta_0 + \beta_1 x + \epsilon$ β_j: 連立 1 次方程式を解いて
カテゴリ変数の確率	$Y \sim Bin(1, p)$ の母比率 p の推定 $\hat{p} = \frac{k}{n}$	ロジスティック回帰 $P(Y=1) = \frac{e^{\boldsymbol{\beta}^\top \cdot \boldsymbol{x}}}{1+e^{\boldsymbol{\beta}^\top \cdot \boldsymbol{x}}}$ β_j: ニュートン法で

ロジスティック回帰で最尤推定

尤度

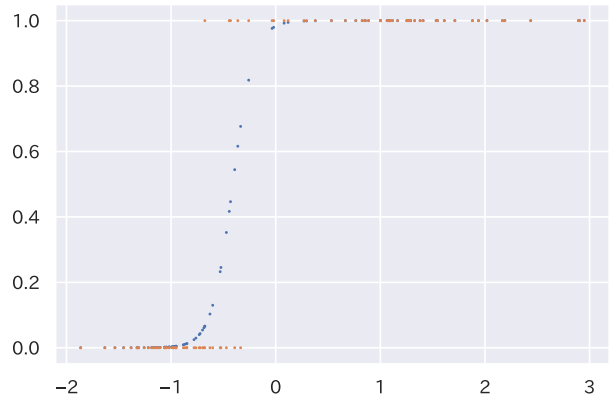
$$L(\boldsymbol{\beta}|y_1,\dots,\boldsymbol{x}_1,\dots) = \prod_{i=1}^{n} p(y_i|\boldsymbol{\beta},\boldsymbol{x}_i) = \prod_{i=1}^{n} (S(\boldsymbol{\beta}^\top \boldsymbol{x}_i))^{y_i} (1 - S(\boldsymbol{\beta}^\top \boldsymbol{x}_i))^{1-y_i}$$

対数尤度

$$\log L(\boldsymbol{\beta}) = \sum_i y_i \log(S(\boldsymbol{\beta}^\top \boldsymbol{x}_i)) + \sum_i (1-y_i) \log(1 - S(\boldsymbol{\beta}^\top \boldsymbol{x}_i)).$$

$\frac{\partial}{\partial \beta_0} \log L = \frac{\partial}{\partial \beta_1} \log L = 0$

$\rightsquigarrow$ ニュートン法数値計算法などで $\hat{\boldsymbol{\beta}}$ を決める.



L06-Q2

Quiz(定数項のみのロジスティック回帰)

ロジスティック回帰の一定モデル $\boldsymbol{\beta} = (\beta_0), \boldsymbol{x} = (1)$ について，データ $(x_i, y_i)_{i=1,\dots,n}$ から β_0 を最尤推定しよう．

ここまで来たよ

6. ロジスティック回帰・分類モデルの評価
   - 単回帰・重回帰では変数選択重要
   - 分類問題
   - ロジスティック回帰の統計モデル
   - 分類問題の精度評価

混同行列

回帰の決定係数 R^2 のようなのりで，予測の正確さを評価したい．
$p(x) \geq c$ のとき $\hat{y} = 1$ と決めることにする．
境い目 c は閾値 (いきち，しきいち, threshold) とよばれる値の例．

2 値分類なので $(\hat{y} - y)^2 = 0, 1$ の区別しかない．
そこで，$(\hat{y}, y) = (0,0), (0,1), (1,0), (1,1)$ で 4 分割して考える．

$y = 0, 1$ は対称だが，以下では，$y = 1$ の方が，特別なデータ点，「検出」したい異常なデータ点，「陽性」(positive) の方だとして区別して考える習慣．$y = 0$ は陰性 (negative)，正常，

実測 y\予測 $\hat{y}$	1	0
1	TP 真陽性	FN 見逃し, 偽陰性
0	FP 誤検出, 偽陽性	TN 真陰性

真/偽=True/False, 陽性/陰性=Postive/Negative

統計的仮説検定, 第 1 種の過誤，第 2 種の過誤, 実測=真の母数，予測=検定の結果 確率統計 I

混同行列 (confusion matrix) 上の度数を表にして行列と思ったもの．

実測 y\予測 $\hat{y}$	1	0
1	TP 真陽性	FN 見逃し, 偽陰性
0	FP 誤検出, 偽陽性	TN 真陰性

真/偽=True/False, 陽性/陰性=Postive/Negative
以下は大きい方がいいが, 一つを大きくすると他が小さくなりがち

- 適合率=精度=precision= $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$
- 感度=再現率=真陽性率=recall=sensitivity= $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$
- 特異度=真陰性率=specificity= $\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$
- 正解率=accuracy $= \frac{\mathrm{TP}+\mathrm{TN}}{\text{すべて}}$
- F 値=F_1 スコア=recall と precision の調和平均 $= \frac{1}{\frac{1}{2}(\frac{1}{\mathrm{recall}}+\frac{1}{\mathrm{precision}})}$

以下は小さい方がいい

- 偽陽性率=1−specificity= $\frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}$

ROC=Curve=Receiver Operating Characteristic

$0 < c < 1$ をどの値にとると，いい感じになるか？

- 激辛 $c = 1$ なら precision, specificity は 1 までよくなる．
- 激甘 $c = 0$ なら recall は 1 までよくなる．

c を動かした時の振る舞い全体を表現するのが ROC という図．

- 縦軸 recall= 真陽性率 $= \frac{\mathrm{TP}}{\mathrm{TP+FN}}$
- 横軸 1−specificity= 偽陽性率 $= \frac{\mathrm{FP}}{\mathrm{TN+FP}}$
- 左上ほどよい

c をパラメタとして動かす．（重複がなければ）テストデータと同じ個数の点が正方形内に打たれる．

- $c = 1$ で左下 (真陽性, 偽陽性)$= (0, 0)$
- $c = 0$ で右上 (真陽性, 偽陽性)$= (1, 1)$

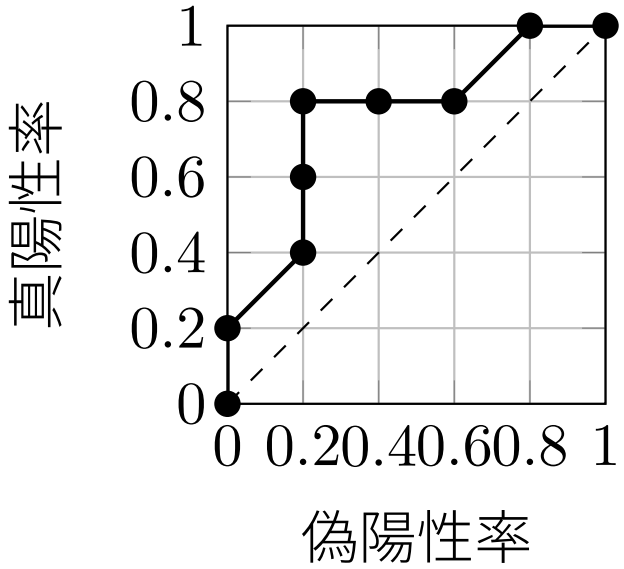
これらの点を線分でつないだ折れ線が ROC「曲線」．
コインを投げるランダム判定なら対角線．それが左上に寄るほど性能がよい．

AUC=Area(面積) under the curve

ROC を 1 個の値 $0 \leq \mathrm{AUC} \leq 1$ で評価できる.
AUC:=ROC の曲線の右下の面積.

- AUC=1 完璧
- AUC=0.5 当てずっぽう
- AUC< 0.5 ◯ × 問題で 50% 未満. 逆に答えた方がましだったのに

ROC



ROC

ロジスティック回帰（または他の手法）により，目的変数の値 y, 説明変数の値 x に対する予測値 $\hat{p}(x)$ を得た．

1. 閾値 $c = 0.3$ とし，$\hat{p}(x) \geq c$ を陽性 $\hat{y} = 1$, それ以外を陰性 $\hat{y} = 0$ と分類するときに，混同行列，再現率 (recall)=真陽性率，適合率 (precision)，偽陽性率，F_1 スコアを求めよう．
2. ROC を描こう．

$\hat{p}(x)$	y
0.9	1
0.8	1
0.8	0
0.7	1
0.6	1
0.4	0
0.3	0
0.2	1
0.2	0
0.1	0