労働経済学
Lecture 2 数学的準備

張　俊超

26th April 2017

ミクロ経済学の復習

- 総和記号、一次関数、二次関数、対数関数
- 微分
- 選好、無差別曲線
- 予算制約条件で効用最大化

総和記号

$\sum_{i=1}^{n} y_i = y_1 + y_2 + \ldots + y_{n-1} + y_n$

一次関数、二次関数

1. **一次関数**
   $y = f(x) = kx + b$

2. **二次関数の最大/最小値**
   $y = f(x) = ax^2 + bx + c$
   $f'(x) = 2ax + b = 0 \Rightarrow x = -\frac{b}{2a}$

対数関数

1. $\ln XY = \ln X + \ln Y$
2. $\ln \frac{X}{Y} = \ln X - \ln Y$
3. $\ln X^Y = Y \ln X$
4. $\ln 1 = 0$
5. $\ln 0$ 定義されていない

微分の公式

1. $c^{'} = 0$
2. $x^{'} = 1$
3. $cx^{'} = c$
4. $(x^c)^{'} = cx^{c-1}$
5. $(\frac{1}{x})^{'} = (x^{-1})^{'} = -x^{-2} = -\frac{1}{x^2}$
6. $(e^x)^{'} = e^x$
7. $(\ln x)^{'} = \frac{1}{x}$

微分の法則

1. $(cf(x))' = cf'(x)$
2. $(f(x) + g(x))' = f'(x) + g'(x)$
3. $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$
4. $\left(\frac{1}{f(x)}\right)' = \frac{-f'(x)}{f^2(x)}$
5. $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$
6. $f'(g(x)) = f'(g(x))g'(x)$
   例：$f(x) = (x^2 + 1)^3$
   $f'(x) = 3(x^2 + 1)^2(2x)$

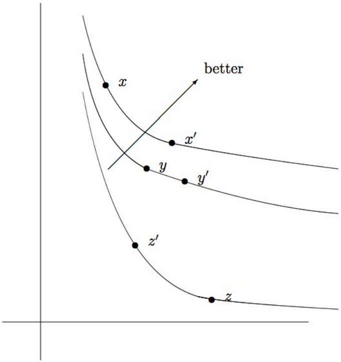
選好

- ミクロ経済学において、財に対する選好は常に以下の仮定をしている。簡単化のために、消費集合 $X = (x, y, z)$ にしする。x, y, z の三種類の財しかない。
  - 反射性：X 集合上の任意の要素 x について、$x \preceq x$ が成り立つ
    例：りんごはりんごより厳密に望ましいか両者は無差別 (りんごはりんごより悪くはない)
  - 推移性：X 集合上の任意の 3 要素 x, y, z について、$x \preceq y$ かつ $y \preceq z$ ならば必ず $x \preceq z$ が成り立つ。
    例：みかんはりんごより悪くはない、かつ、すいかはみかんより悪くはない場合、すいかは必ずりんごより悪くはない。
  - 完全性：X 集合上の任意の 2 要素 x, y について、必ず $x \preceq y$ もしくは $y \preceq x$ のいずれか少なくとも一方が成り立つ。
    例：みかんはりんごより悪くはなければ、必ずりんごはみかんより悪くはない。
  - 選好順序＝反射性 + 完備性 + 推移性 $\Rightarrow$ 合理的な選好

効用関数

- 経済学では、抽象的な選好を表すために、効用関数を用いる。効用関数は労働者の間で、比較できません。
- 一人の労働者にとって、選好 $\Longleftrightarrow$ 効用関数。
- 任意の二つの要素の (数量の) 組み合わせを考えましょう。(x_a, y_a) の組み合わせを A、(x_b, y_b) を B とする。労働者が A を B よりも好む時、$U(x_a, y_a) > U(x_b, y_b)$。B を A よりも好む時、$U(x_a, y_a) < U(x_b, y_b)$。A と B に対する選好は等しい時、$U(x_a, y_a) = U(x_b, y_b)$。このような U が効用関数。

無差別曲線



- 効用関数の 2 財平面上の等高線図を無差別曲線と呼ぶ。
- 無差別曲線は一般的に凸性を持つ。曲線は原点に向くって、凸であること。
- 各曲線上、効用は同じ。しかし、原点から遠ければ遠いほど、効用が高くなる。

効用最大化

$$
\begin{aligned}
&Max \ \ U(x,y) \\
s.t. \ \ & p_x x + p_y y = I
\end{aligned}
\tag{1}
$$

- p_x, p_y は財 x,y の値段、I は所得。
- 効用 U を最大化するような、財な数量、x と y を決める。
- (1) 式の解け方：ラグランジェ乗数法

ラグランジェ乗数法

1. 最初に、目的関数を書く
   $U(x, y)$
2. そして、とにかく λ という不思議の数を隣に書く
   $U(x, y) + \lambda$
3. 予算制約 $p_x x + p_y y = I$ の右辺を固定し、固定していない左辺を引いたものを λ の隣に書く
   $U(x, y) + \lambda(I - (p_x x + p_y y))$
4. この関数の変数は x, y, λ となり、ラグランジェ関数と呼ぶ。
   $\mathcal{L}(x, y, \lambda) = U(x, y) + \lambda(I - (p_x x + p_y y))$
5. ラグランジェ関数をそれぞれ x, y, λ について偏微分して、連立方程式を解ける。
   $\frac{\partial \mathcal{L}(x,y,\lambda)}{\partial x} = \frac{\partial U(x,y)}{\partial x} - \lambda p_x = 0$
   $\frac{\partial \mathcal{L}(x,y,\lambda)}{\partial y} = \frac{\partial U(x,y)}{\partial x} - \lambda p_y = 0$
   $\frac{\partial \mathcal{L}(x,y,\lambda)}{\partial \lambda} = I - (p_x x + p_y y) = 0$

数理統計学の復習

- 確率変数、確率分布、分布関数
- 期待値と分散
- 共分散と相関

確率変数、確率分布、分布関数

- $x_1, x_2, ..., x_n$ なる n 個の値をとる変数 X に対して、$X = x_i$ なる確率 p_i が与えられているとき、X を確率変数という。
- 確率変数 X とそれに対応する確率 $P(X = x_i)$ との対応関係を確率分布という。
- 確率変数 X の値がある値 x までとる確率を $F(x)$ で表し，確率変数 X の分布関数という。つまり，分布関数は $F(x) = P(X \leq x)$ で与えられる。

離散型

Xの値 x_i	x_1	x_2	$\cdots$	x_n
$P(X=x_i)=p_i=f(x_i)$	p_1	p_2	$\cdots$	p_n

- 確率分布は上記の表どおり。
- 確率変数 X の値を $x_1 < x_2 < ...x_n$ とするときに、分布関数 $F(x_k) = P(X \leq x_k) = p_1 + p_2 + p_3 + ... + p_k = \sum_{i=1}^{k} p_i$
- 確率分布 f と分布関数 F は次の性質をもつ。

1. $0 \leq p_i = f(x_i) \leq 1$
2. $F(x_k) = P(X \leq x_k) = p_1 + p_2 + p_3 + ... + p_k = 1$
3. $P(a < X \leq b) = F(b) - F(a)$
4. $a < b \Longleftrightarrow F(a) < F(b)$

連続型

- 確率変数が連続的な値であれば、分布関数は以下で $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$
- $f(x)$ は確率密度関数という。以下の性質を持つ。

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. 任意の実数 $x_1, x_2 (x_1 < x_2)$ について、$P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\,dx$
4. $f(x)$ は x において連続であれば、$F^{'}(x) = f(x)$

期待値と分散

- **離散型** $p_i = P(X = i)$
  $E(X) = x_1p_1 + x_2p_2 + \ldots + x_kp_k = \sum_{i=1}^{k} x_ip_i$
  $V(X) = (x_1 - E(X))^2p_1 + (x_2 - E(X))^2p_2 + \ldots = \sum_{i=1}^{k}(x_i - E(X))^2p_i$
- **連続型** $P(X \leq x) = F(x) = \int_{-\infty}^{x} f(x)\,dx$
  $E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$
  $V(X) = \int_{-\infty}^{\infty}(x - E(X))^2f(x)\,dx$
- **演算のルール**
  $E(aX + b) = aE(X) + b$
  $V(aX + b) = a^2\,V(X)$

共分散と相関

- 共分散 Cov(X,Y) は X と Y との間の関係を表す数値である。
  $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$
  共分散が大きい（正）$\rightarrow$ X が大きいとき Y も大きい傾向がある
  共分散が 0 に近い $\rightarrow$ X と Y にあまり関係はない
  共分散が小さい（負）$\rightarrow$ X が大きいとき Y は小さい傾向がある

- 相関係数
  $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$
  ρ_{XY} は-1 と 1 との間にある。$|\rho_{XY}| = 1$ の場合、完全相関であり、X と Y は線形関係である。$\rho_{XY} = 0$ の場合、相関しない、全く線形関係ではない。

計量経済学の復習

- 労働経済学における実証分析に最小限の数学
- 計量経済学の基礎となる最小二乗法を復習
- 単回帰と重回帰
- OLS における仮定と BLUE 推定量

最小二乗法 (Ordinary Least Squares, OLS)

- なぜ必要か？

  感覚的になんとなくわかることを数学的に表現できるように

  (定性的 $\longrightarrow$ 定量的)

- どんな時使うか？

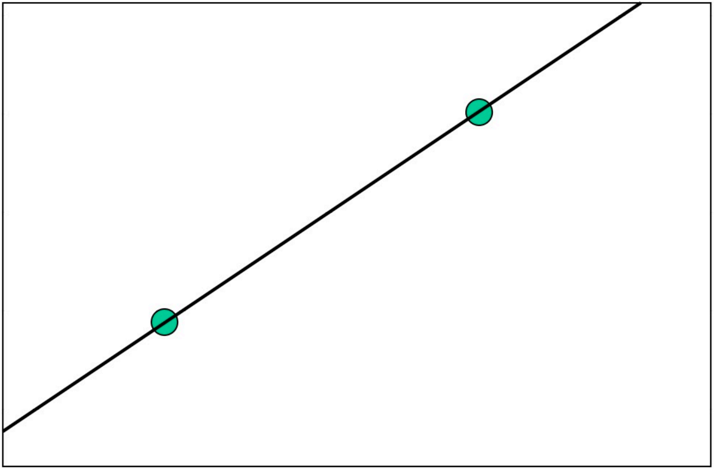
  あいまいな情報を数字で表現したい時 (トレンド分析、変数の関係を分析...)

  データ解析を定量化したい時 (データの解釈と予測、データの近似式、データからの最適設計)

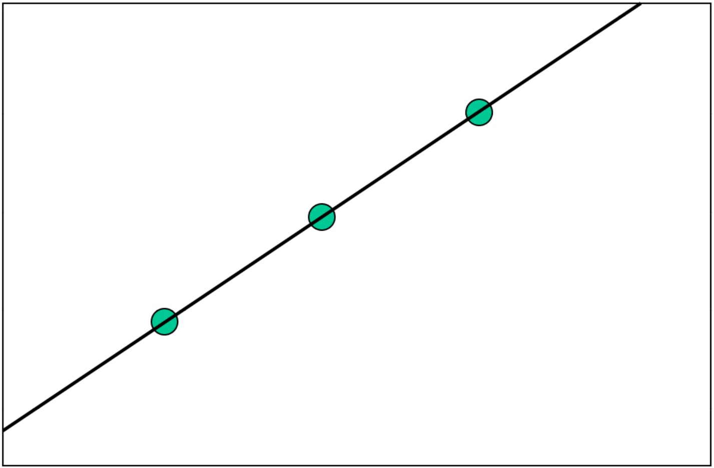
- 本質となる考え方

  統計的に「誤差」を定義し、もっともらしい直線を引く

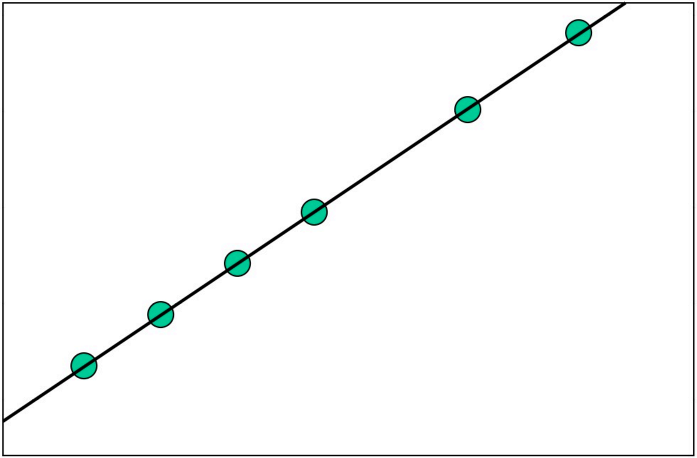
最小二乗法のイメージ (直線当てはめ)



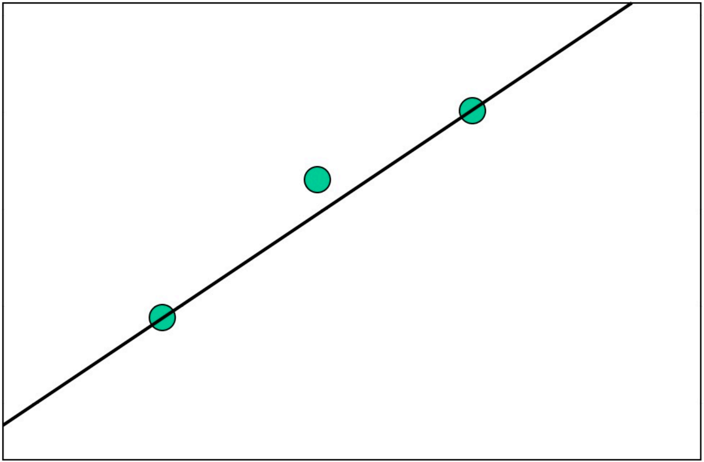
最小二乗法のイメージ (直線当てはめ)



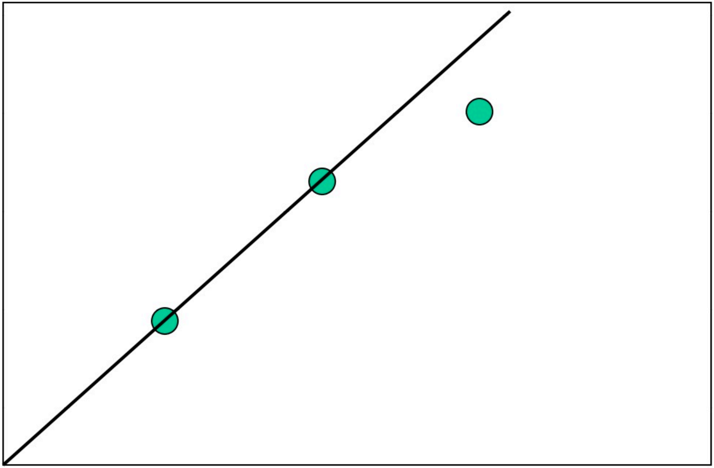
最小二乗法のイメージ (直線当てはめ)



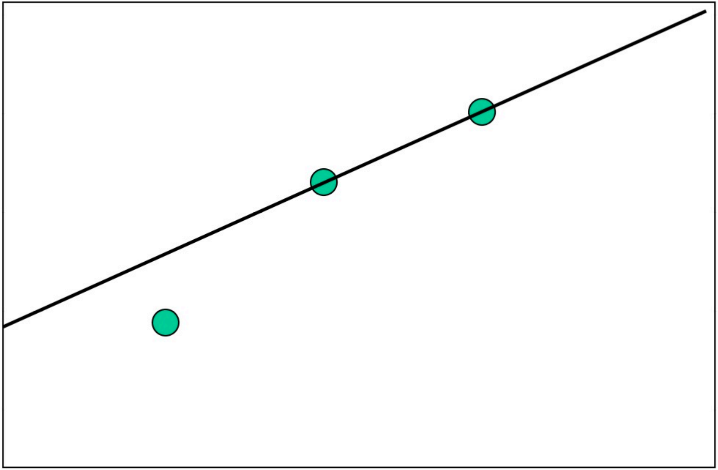
最小二乗法のイメージ (直線当てはめ)



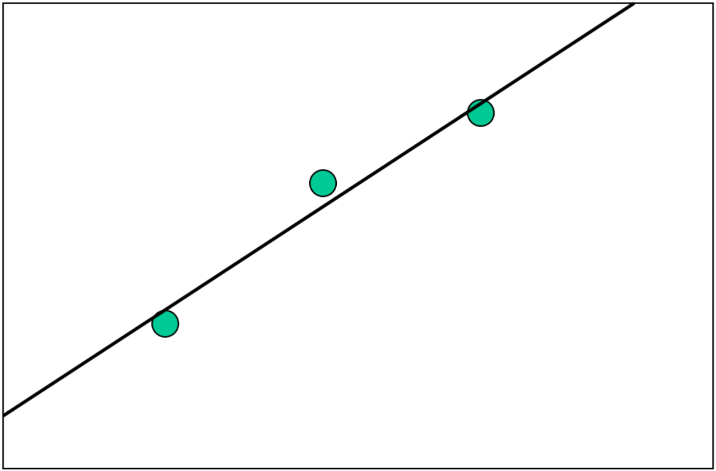
最小二乗法のイメージ (直線当てはめ)



最小二乗法のイメージ (直線当てはめ)

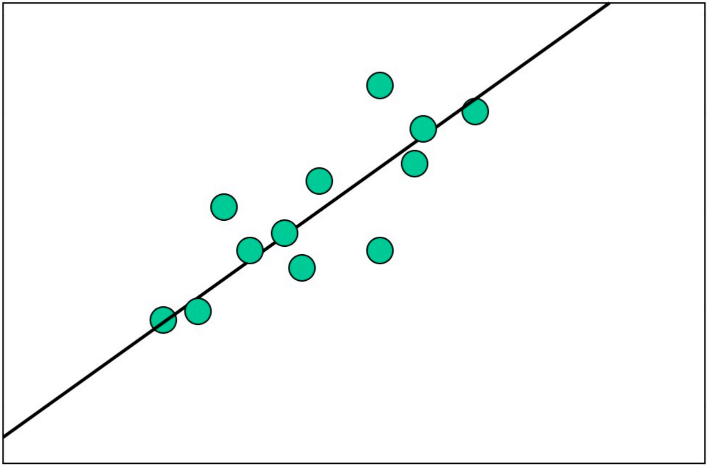


最小二乗法のイメージ (直線当てはめ)



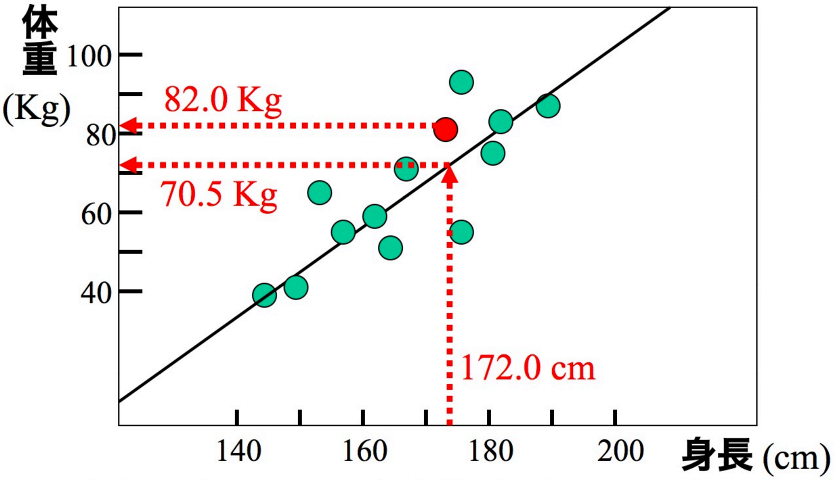
点が三つ以上になれば、直線当てはめには何らかの妥協が必要

最小二乗法のイメージ (直線当てはめ)



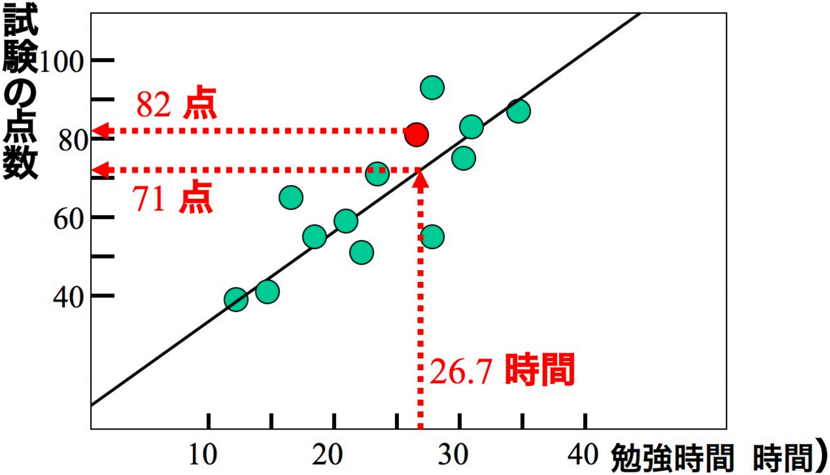
点が多くなると、直線当てはめには数学的基準が必要

最小二乗法の応用例



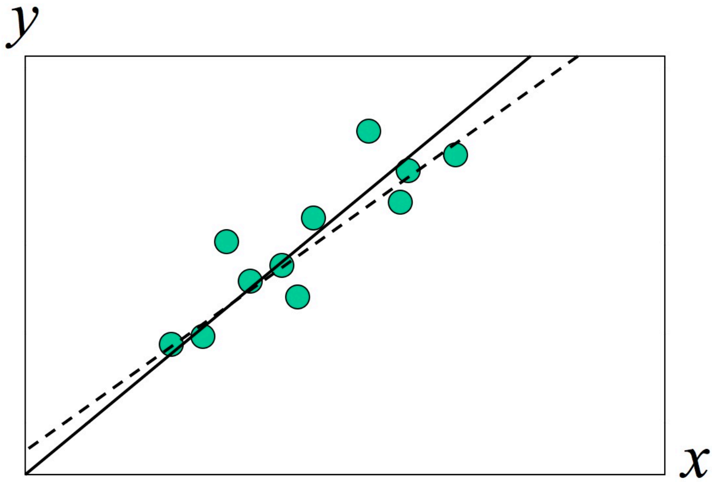
自分の身長から、平均的体重を推測できる

最小二乗法の応用例



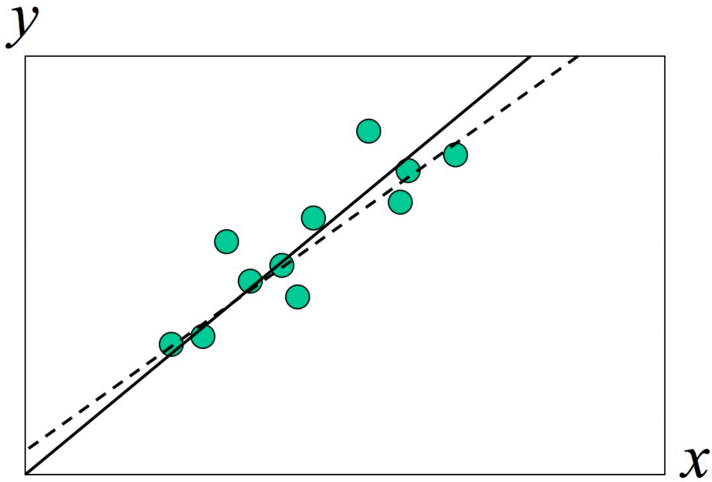
長い時間勉強をすれば、一般的に試験で良い点数が取れる

最小二乗法の統計的基準



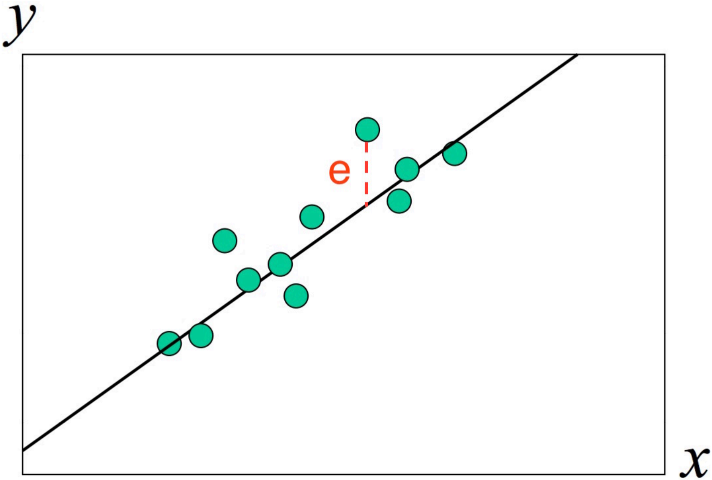
多数のデータから、直線の数式表現 ($y = ax + b$) を求める
(x,y) はデータから観察される

最小二乗法の統計的基準



多数のデータから、直線の数式表現 ($y = ax + b$) を求める
(x,y) はデータから観察される

最小二乗法の統計的基準



残差 e は垂直方向において各点から直線までの距離と定義される。(仮に $y = ax + b$ の関数型がわかる場合)

最小二乗法の統計的基準

n 個の点があり、i 番目の点の座標値を (x_i, y_i)

$e_i = y_i - (ax_i + b)$

RHS と LHS を両方二乗にして、

$e_i^2 = \{y_i - (ax_i + b)\}^2$

n 個の点の e_i^2 の和を取れば、

$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\{y_i - (ax_i + b)\}^2 \equiv S$

S は残差二乗和と定義され、SSR または RSS という表記もある

最小二乗法の統計的基準

S を最小化することによって、a と b の値が得られる

$arg\ min\ S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\{y_i - (ax_i + b)\}^2$

$$\begin{cases} \frac{\partial S}{\partial a} = 0 \\ \frac{\partial S}{\partial b} = 0 \end{cases} \tag{2}$$

$\Longrightarrow$

$$\begin{cases} \frac{\partial S}{\partial a} = \frac{\partial \sum_{i=1}^{n}\{y_i - (ax_i + b)\}^2}{\partial a} = \sum_{i=1}^{n} -2(y_i - ax_i - b)x_i = 0 \\ \frac{\partial S}{\partial b} = \frac{\partial \sum_{i=1}^{n}\{y_i - (ax_i + b)\}^2}{\partial b} = \sum_{i=1}^{n} -2(y_i - ax_i - b) = 0 \end{cases} \tag{3}$$

最小二乗法の統計的基準

$\Longleftrightarrow$

$$\begin{cases} n^{-1}\sum_{i=1}^{n}(y_i - ax_i - b)x_i = 0 \\ n^{-1}\sum_{i=1}^{n}(y_i - ax_i - b) = 0 \end{cases} \tag{4}$$

$\sum_{i=1}^{n} x_i(x_i - \bar{x}) = \sum_{i=1}^{n}(x_i - \bar{x})^2$ と $\sum_{i=1}^{n} x_i(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ の性質を使って、

$\Longleftrightarrow$

$$\begin{cases} a = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\ b = \bar{y} - a\bar{x} \end{cases} \tag{5}$$

よって、a と b を得られる。その他、行列の式、モーメント法、最大尤度法でも a と b を推定できる。違う手法で a と b を計算してみてください。

単回帰と重回帰

- **単回帰**

  $y_i = a + bx_i + e_i$

  左辺 y_i は被説明変数。従属変数、または結果変数とも呼ばれる。

  右辺の x_i は説明変数。独立変数、または原因変数とも呼ばれる。

  説明変数は一つだけ。

- **重回帰**

  $y_i = a + bx_i + cz_i + ... + e_i$

  説明変数：$x_i, z_i, ...$

  説明変数が二つ以上。

単回帰と重回帰

- (4) 式では、単回帰での a, b を計算した。$S = f(a, b)$ を最小化するような a,b を得られた。
- 重回帰の場合、同様に、$S = f(a, b, c, ...)$ を最小化するような a,b,c,... を得られる。

$$\frac{\partial S}{\partial a} = 0$$

$$\frac{\partial S}{\partial b} = 0$$

$$\frac{\partial S}{\partial c} = 0$$

...

最小二乗法 (OLS) における仮定

1. 線形性
   真のモデルは $y_i = \alpha + \beta x_i + u_i$ に従う。α, β は推定すべき未知パラメータ。u_i は誤差の実現値。
2. 誤差項の期待値は 0
   $E(u_i) = 0$
3. 誤差項の分散はすべての i について等しい
   $Var(u_i) = \sigma^2$
4. 誤差項に系列相関は存在しない
   $Cov(u_i, u_j) = 0$
5. 説明変数と誤差項の独立性
   $Cov(u_i, x_i) = 0$
6. 正規分布の仮定
   $u_i \sim N(0, \sigma^2)\ i.i.d.$

仮定 1-5 を満たす場合、OLS 推定量は BLUE(Best Linear Unbiased Estimator) である。最良線形不偏推定量である。

決定係数 (R^2)、相関係数、t 検定、F 検定、p 値

- 計量経済学を履修してください。
- または、Wooldridge のテキストを独学してください。
- それらの知識がわからなくても講義内容の理解には問題ない。