最小自乗法(OLS)

研究演習 I
神戸大学・経済学部

1 回帰モデル

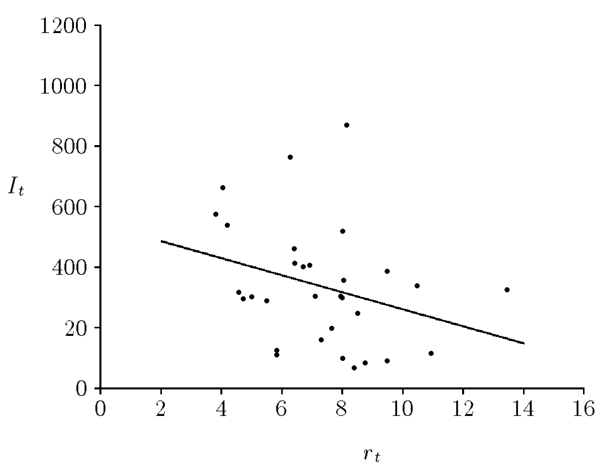
(x_1, y_1), (x_2, y_2), $\cdots$, (x_T, y_T) のように T 組のデータがあり，x_t と y_t との間に以下の線型関係を想定する。

$$y_t = \alpha + \beta x_t + u_t, \tag{1}$$

ただし，u_t は誤差項，または，撹乱項と呼ばれる。x_t は説明変数，y_t は被説明変数，α, β はパラメータとそれぞれ呼ばれる。

上の式は回帰モデル (または，回帰式) と呼ばれる。目的は，切片 α と傾き β をデータ $\{(x_t, y_t), t = 1, 2, \cdots, T\}$ から推定すること，すなわち，図 1 において直線を求めることにある。

図 1： 投資 (y_t) と利子率 (x_t)



2 撹乱項 u_t の仮定

撹乱項 u_1, u_2, $\cdots$, u_T はそれぞれ独立に平均ゼロ，分散 σ^2 の正規分布する。すなわち，

$$u_t \sim N(0, \sigma^2) \tag{2}$$

$$\mathrm{E}(u_t u_s) = \begin{cases} \sigma^2, & t = s \text{ のとき}, \\ 0, & t \neq s \text{ のとき}, \end{cases} \tag{3}$$

3 切片 α と傾き β の推定

次のような関数 $S(\alpha, \beta)$ を定義する。

$$S(\alpha, \beta) = \sum_{t=1}^{T} (y_t - \alpha - \beta x_t)^2 \tag{4}$$

このよき，

$$\min_{\alpha, \beta} S(\alpha, \beta) \tag{5}$$

となるような α, β を求める (最小自乗法)。このときの解を $\widehat{\alpha}$, $\widehat{\beta}$ とする。

最小化のためには，

$$\begin{aligned} \frac{\partial S(\alpha, \beta)}{\partial \alpha} &= 0 \\ \frac{\partial S(\alpha, \beta)}{\partial \beta} &= 0 \end{aligned} \tag{6}$$

を満たす α, β が $\widehat{\alpha}$, $\widehat{\beta}$ となる。

二つの連立方程式を解くと，

$$\begin{aligned} \widehat{\beta} &= \frac{\sum_{t=1}^{T} (x_t - \overline{x})(y_t - \overline{y})}{\sum_{t=1}^{T} (x_t - \overline{x})^2} \\ \widehat{\alpha} &= \overline{y} - \widehat{\beta}\overline{x} \end{aligned} \tag{7}$$

となる。ただし，

$$\overline{x}=\frac{1}{T}\sum_{t=1}^{T}x_t,\quad \overline{y}=\frac{1}{T}\sum_{t=1}^{T}y_t, \tag{8}$$

とする。

4　多変数の回帰モデル

T 組のデータが $y_t, x_{1t}, x_{2t}, \cdots, x_{kt})$, $t=1,2,\cdots,T$, であり，その関係が

$$y_t=\beta_1 x_{1t}+\beta_2 x_{2t}+\ \cdots\ \beta_k x_{kt}+u_t \tag{9}$$

となる場合を考える (すべての t について，$x_{1t}=1$ とすれば，β_1 は切片となる)。y_t は被説明変数，x_{1t}, x_{2t}, $\cdots$, x_{kt} は説明変数，β_1, β_2, $\cdots$, β_k はパラメータ，u_t は攪乱項とそれぞれ呼ばれる。

$$\begin{aligned}&S(\beta_1,\beta_2,\cdots,\beta_k)\\&=\sum_{t=1}^{T}(y_t-\beta_1 x_{1t}-\beta_2 x_{2t}-\ \cdots\ -\beta_k x_{kt})^2\end{aligned} \tag{10}$$

$S(\beta_1,\beta_2,\cdots,\beta_k)$ を最小にするような $\beta_1,\beta_2,\cdots,\beta_k$ を求める。その解を $\widehat{\beta}_1,\widehat{\beta}_2,\cdots,\widehat{\beta}_k$ とする。

5　攪乱項の分散 σ^2 の推定

攪乱項の分散 σ^2 の推定値を s^2 とする。

$$s^2=\frac{1}{T-k}\sum_{t=1}^{T}(y_t-\widehat{\beta}_1 x_{1t}-\ \cdots\ -\widehat{\beta}_k x_{kt})^2 \tag{11}$$

$T-k$ は自由度と呼ばれる (この場合，推定するパラメータは $\beta_1,\beta_2,\cdots,\beta_k$ の k 個なので，T から k を引く)。

s は，回帰式の標準誤差 (Standard Error of Regression) と呼ばれる。

6　$\widehat{\beta}_i$ の分布

$\widehat{\beta}_i$, $i=1,2,\cdots,k$, の分布は以下の通り (分布の導出は省略)。

$$\frac{\widehat{\beta}_i-\beta}{s_{\beta_i}}\ \sim\ t(T-k) \tag{12}$$

となる (自由度 $T-k$ の t 分布)。ただし，s_{β_i} は $\widehat{\beta}_i$ の標準誤差 (Standard Error) と呼ばれる。

7　仮説 $H_0:\beta_i=0$ の検定 (t 値)

仮説 H_0 : $\beta_i=0$ が正しいとするとき，

$$t_0=\frac{\widehat{\beta}_i}{s_{\beta_i}}\ \sim\ t(T-2) \tag{13}$$

となる (自由度 $T-2$ の t 分布)。

t_0 を t 値 (t-value, または，t-statistic) と呼ぶ。

t_0 が絶対値で大きければ，t_0 は自由度 $T-2$ の t 分布の端にあり，起こり得る可能性が低いということを意味する。t_0 が絶対値で大きいかどうかの判定には，t 分布表が必要。例えば，豊田編『基本統計学』(p.171, 付表3) を参照せよ。

8　決定係数 R^2 について

回帰式の当てはまりの良さを示す指標として，次の式で定義される決定係数 (R-squared) が用いられる。

$$R^2=1-\frac{\sum_{t=1}^{T}\widehat{u}_t^2}{\sum_{t=1}^{T}(y_t-\overline{y})^2}, \tag{14}$$

ただし，$\widehat{u}_t$ は残差と呼ばれ，

$$\widehat{u}_t=y_t-\widehat{\beta}_1 x_{1t}-\widehat{\beta}_2 x_{2t}-\ \cdots\ -\widehat{\beta}_k x_{kt}, \tag{15}$$

として定義される (攪乱項 u_t の推定値とも考えられる)。

明らかに，

$$0\le R^2\le 1, \tag{16}$$

となる。R^2 が 1 に近づけば回帰式の当てはまりは良いと言える。しかし，t 分布のような数表は存在しない。したがって,「どの値よりも大きくなるべき」というような基準はない。

慣習的には，メドとして 0.9 以上を判断基準にする。

表 1: t 分布表 – $t(m)$

$$\alpha = \mathrm{Prob}(t > t_\alpha) = \int_{t_\alpha}^{\infty} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \frac{1}{\sqrt{m\pi}} \frac{1}{(1+\frac{x^2}{m})^{\frac{m+1}{2}}} \mathrm{d}x$$

α / m	.10	.05	.025	.001	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
31	1.309	1.696	2.040	2.453	2.744
32	1.309	1.694	2.037	2.449	2.738
33	1.308	1.692	2.035	2.445	2.733
34	1.307	1.691	2.032	2.441	2.728
35	1.306	1.690	2.030	2.438	2.724
36	1.306	1.688	2.028	2.434	2.719
37	1.305	1.687	2.026	2.431	2.715
38	1.304	1.686	2.024	2.429	2.712
39	1.304	1.685	2.023	2.426	2.708
40	1.303	1.684	2.021	2.423	2.704
41	1.303	1.683	2.020	2.421	2.701
42	1.302	1.682	2.018	2.418	2.698
43	1.302	1.681	2.017	2.416	2.695
44	1.301	1.680	2.015	2.414	2.692
45	1.301	1.679	2.014	2.412	2.690
46	1.300	1.679	2.013	2.410	2.687
47	1.300	1.678	2.012	2.408	2.685
48	1.299	1.677	2.011	2.407	2.682
49	1.299	1.677	2.010	2.405	2.680
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
80	1.292	1.664	1.990	2.374	2.639
120	1.289	1.658	1.980	2.358	2.617
240	1.285	1.651	1.970	2.342	2.596
∞	1.282	1.645	1.960	2.326	2.576

9 自由度修正済み決定係数 $\overline{R}^2$ について

決定係数 R^2 によると，説明変数を増やせば増やすほど R^2 は大きくなる。y_t にとっては意味の無い変数でも，回帰式に加えれば R^2 は大きくなる。これを回避するために，回帰式の当てはまりの良さを示す指標に自由度修正済み決定係数 (Adjusted R-squared) を用いる。定義は以下の通りである。

$$\overline{R}^2 = 1 - \frac{\sum_{t=1}^{T} \widehat{u}_t^2/(T-k)}{\sum_{t=1}^{T}(y_t-\overline{y})^2/(T-1)}, \tag{17}$$

決定係数と自由度修正済み決定係数との間には，$\overline{R}^2 \leq R^2$ という関係が成り立つ。

10 ダービン・ワトソン比 (DW) について

最小自乗法の仮定の一つに,「攪乱項 $u_1, u_2, \cdots, u_T$ はそれぞれ独立に分布する」というものがあった。ダービン・ワトソン比 (Durbin-Watson statistic) とは，誤差項の系列相関，すなわち，u_t と u_{t-1} との間の相関の有無を検定するために考案された。

すなわち，ダービン・ワトソン比とは，回帰式が

$$y_t = \alpha + \beta x_t + u_t,$$
$$u_t = \rho u_{t-1} + \epsilon_t,$$

のときに，$H_0:\ \rho = 0$, $H_1:\ \rho \neq 0$ の検定である。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ は互いに独立とする。

ダービン・ワトソン比の定義は次の通りである。

$$DW = \frac{\sum_{t=2}^{T}(\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=1}^{T}\widehat{u}_t^2} \tag{18}$$

DW は近似的に，次のように表される (証明は略)。

$$DW \approx 2(1-\widehat{\rho}), \tag{19}$$

$\widehat{\rho}$ は $\widehat{u}_t$ と $\widehat{u}_{t-1}$ の相関係数

1. DW の値が 2 前後のとき，系列相関なし ($\widehat{\rho} = 0$ のとき，$DW \approx 0$)。

2. DW が 2 より十分に小さいとき，正の系列相関

3. DW が 2 より十分に大きいとき，負の系列相関

正確な判定には，データ数 T とパラメータ数 k に依存する (数表では，データ数は n を用いている)。表 2 と表 3 を参照せよ。

図 2： 正の系列相関

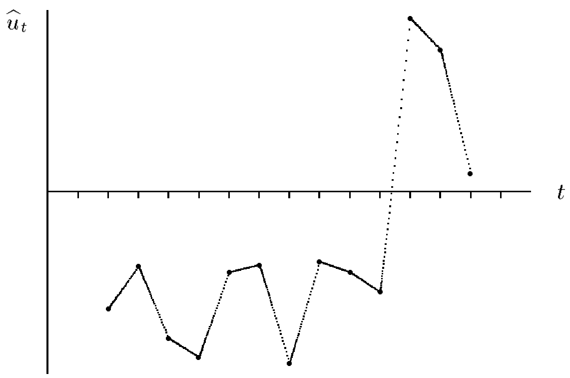


図 3： 負の系列相関

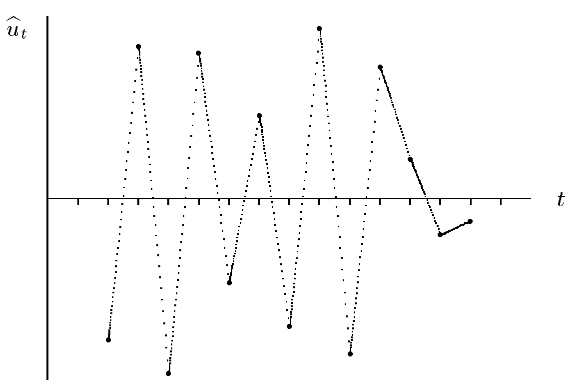


表 2: ダービン・ワトソン統計量の 5 % 点の上限と下限

(1) $k=2$

T	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	1.08	1.08	1.36	1.36	2.64	2.64	2.92	2.92	4
20	0	1.20	1.20	1.41	1.41	2.59	2.59	2.80	2.80	4
25	0	1.29	1.29	1.45	1.45	2.55	2.55	2.71	2.71	4
30	0	1.35	1.35	1.49	1.49	2.51	2.51	2.65	2.65	4

(2) $k=3$

T	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.95	0.95	1.54	1.54	2.46	2.46	3.05	3.05	4
20	0	1.10	1.10	1.54	1.54	2.46	2.46	2.90	2.90	4
25	0	1.21	1.21	1.55	1.55	2.45	2.45	2.79	2.79	4
30	0	1.28	1.28	1.57	1.57	2.43	2.43	2.72	2.72	4

(3) $k=4$

T	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.82	0.82	1.75	1.75	2.25	2.25	2.25	3.18	4
20	0	1.00	1.00	1.68	1.68	2.32	2.32	2.32	3.00	4
25	0	1.12	1.12	1.66	1.66	2.34	2.34	2.34	2.88	4
30	0	1.21	1.21	1.65	1.65	2.35	2.35	2.35	2.79	4

(4) $k=5$

T	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.69	0.69	1.97	1.97	2.03	2.03	3.31	3.31	4
20	0	0.90	0.90	1.83	1.83	2.17	2.17	3.10	3.10	4
25	0	1.04	1.04	1.77	1.77	2.23	2.23	2.96	2.96	4
30	0	1.14	1.14	1.74	1.74	2.26	2.26	2.86	2.86	4

(5) $k=6$

T	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.56	0.56	2.21	—	—	2.21	3.44	3.44	4
20	0	0.79	0.79	1.99	1.99	2.01	2.01	3.21	3.21	4
25	0	0.95	0.95	1.89	1.89	2.11	2.11	3.05	3.05	4
30	0	1.07	1.07	1.83	1.83	2.17	2.17	2.93	2.93	4

A: 正の系列相関あり
B: 系列相関の有無を判定不能
C: 系列相関なし
D: 系列相関の有無を判定不能
E: 負の系列相関あり

表 3: ダービン・ワトソン統計量の 5 % 点の上限と下限

	$k=2$		$k=3$		$k=4$		$k=5$		$k=6$	
T	dl	du	dl	du	dl	du	dl	du	dl	du
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

11　ダミー変数について

データに異常値が含まれている場合，経済構造がある時期から変化した場合，ダミー変数を使う。

ダミー変数とは，0 と 1 から成る変数のことである。

例えば，データが 20 期間あるとして，9 期目のデータが，回帰直線から離れている場合 (異常値の場合) を考える。

$$d_t = \begin{cases} 0, & t \neq 9 \text{ のとき} \\ 1, & t = 9 \text{ のとき} \end{cases}$$

という変数を作り，

$$y_t = \alpha + \delta d_t + \beta x_t + u_t$$

を推定する。δ の推定値 $\widehat{\delta}$ の有意性を調べることによって，異常値かどうかの検定ができる。

次に，9 期目以前と以降とで，経済構造が変化している場合を考える。

$$d_t = \begin{cases} 0, & t = 1, 2, \cdots, 8 \text{ のとき} \\ 1, & 9, 10, \cdots, 20 \text{ のとき} \end{cases}$$

という変数を作り，

$$y_t = \alpha + \delta d_t + \beta x_t + u_t$$

を推定する (定数項だけが変化したと考えた場合)。または，

$$y_t = \alpha + \delta d_t + \beta x_t + \gamma d_t x_t + u_t$$

を推定する (定数項も係数も変化)。

δ や γ の推定値の有意性を調べることによって，構造変化の検定を行うことができる。

12　系列相関：ρ を含む回帰式の推定方法

回帰式が

$$y_t = \alpha + \beta x_t + u_t,$$

$$u_t = \rho u_{t-1} + \epsilon_t,$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ は互いに独立とする。

u_t を消去すると，

$$(y_t - \rho y_{t-1}) = \alpha(1-\rho) + \beta(x_t - \rho x_{t-1}) + \epsilon_t$$

となり，$(y_t - \rho y_{t-1})$, $(x_t - \rho x_{t-1})$ を新たな変数として，最小二乗法を適用する。$\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ は互いに独立とするなので，最小二乗法を適用が可能となる。

ρ の求め方について

1. ダービン・ワトソン比がゼロに近い場合，$\rho = 0$ と近似して，

$$\Delta y_t = \alpha + \beta \Delta x_t + \epsilon_t$$

として推定する (定数項を含める方が多い)。ただし，$\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$ とする。

2. DW は近似的に $DW \approx 2(1-\widehat{\rho})$ と表されるので，DW から ρ の推定値 $\widehat{\rho}$ を逆算して，$(y_t - \widehat{\rho} y_{t-1})$, $(x_t - \widehat{\rho} x_{t-1})$ を新たな変数として，最小二乗法を適用する。

3. 収束計算によって求める。

   (i) $y_t = \alpha_t + \beta x_t + u_t$ に最小二乗法を適用し，$\widehat{u}_1$, $\widehat{u}_2, \cdots, \widehat{u}_T$ を求める。

   (ii) $\widehat{u}_t = \rho \widehat{u}_{t-1} + \epsilon_t$ に最小二乗法を適用し，$\widehat{\rho}$ を求める。

   (iii) $(y_t - \widehat{\rho} y_{t-1})$, $(x_t - \widehat{\rho} x_{t-1})$ を新たな変数として，$(y_t - \widehat{\rho} y_{t-1}) = \alpha^* + \beta(x_t - \widehat{\rho} x_{t-1}) + \epsilon_t$ を計算して，$\widehat{\alpha}$, $\widehat{\beta}$ を求める。ただし，$\widehat{\alpha}^* = \widehat{\alpha}(1-\widehat{\rho})$ として $\widehat{\alpha}$ を計算する。

   (iv) $\widehat{u}_t = y_t - \widehat{\alpha}_t - \widehat{\beta} x_t$ から $\widehat{u}_t$ を再計算して，最小二乗法により $\widehat{\rho}$ を求める。

   (v) (iii) と (iv) を収束するまで繰り返す。

   (iv) で得られた $\widehat{\rho}$ の有意性から，$H_0:\ \rho = 0$ の検定を行う方法もある。

13　多重共線性

回帰式が

$$y_t = \alpha w_t + \beta x_t + u_t$$

の場合を考える。w_t, x_t が外生変数，y_t は内生変数，u_t は互いに独立な撹乱項とする。$w_t = 1$ のとき，α は定数項となる。

w_t と x_t の相関が大きいことを多重共線性が強いと言う。

w_t と x_t の相関が大きい場合は，α, β の推定値は不安定になる。

極端な場合，w_t と x_t の相関が 1 の場合 (完全相関の場合) は，すべての t について，$x_t = w_t$ となる。この場合，回帰式は

$$\begin{aligned} y_t &= \alpha w_t + \beta x_t + u_t \\ &= (\alpha + \beta) x_t + u_t \end{aligned}$$

となり，$\alpha + \beta$ を推定することは可能だが，α, β を別々に推定することはできなくなる。$y_t = \alpha w_t + \beta x_t + u_t$ を推定した場合，$\alpha + \beta$ の推定値が一定値となる $\widehat{\alpha}, \widehat{\beta}$ の組み合わせは無数に存在する。この意味で，$\widehat{\alpha}, \widehat{\beta}$ は不安定であると言える。

厳密には，最小二乗法によると，

$$\sum_{t=1}^{T} u_t^2 = \sum_{t=1}^{T} (y_t - \alpha w_t - \beta x_t)^2$$

を最小にする α, β をその推定値 $\widehat{\alpha}, \widehat{\beta}$ とする。

すなわち，

$$\frac{\partial \sum u_t^2}{\partial \alpha} = -2 \sum_{t=1}^{T} (y_t - \alpha w_t - \beta x_t) w_t = 0$$

$$\frac{\partial \sum u_t^2}{\partial \beta} = -2 \sum_{t=1}^{T} (y_t - \alpha w_t - \beta x_t) x_t = 0$$

の連立方程式を解くことになる。

$$\begin{aligned} \sum_{t=1}^{T} y_t w_t - \widehat{\alpha} \sum_{t=1}^{T} w_t^2 - \widehat{\beta} \sum_{t=1}^{T} x_t w_t = 0 \\ \sum_{t=1}^{T} y_t x_t - \widehat{\alpha} \sum_{t=1}^{T} w_t x_t - \widehat{\beta} \sum_{t=1}^{T} x_t^2 = 0 \end{aligned}$$

行列表示により，

$$\begin{pmatrix} \sum_t y_t w_t \\ \sum_t y_t x_t \end{pmatrix} = \begin{pmatrix} \sum_t w_t^2 & \sum_t x_t w_t \\ \sum_t w_t x_t & \sum_t x_t^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix}$$

$\widehat{\alpha}$, $\widehat{\beta}$ について表すと，

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \sum_t w_t^2 & \sum_t x_t w_t \\ \sum_t w_t x_t & \sum_t x_t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_t y_t w_t \\ \sum_t y_t x_t \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} \begin{pmatrix} d \\ e \end{pmatrix}$$

逆行列を計算して，

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \frac{1}{ac-b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix}$$

完全な多重共線性の場合 ($x_t = w_t$ の場合)，$ac - b^2 = (\sum_t w_t^2)(\sum_t x_t^2) - (\sum_t w_t x_t) = 0$ となる。

また，

$$\mathrm{Var}\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \sigma^2 \begin{pmatrix} \sum_t w_t^2 & \sum_t x_t w_t \\ \sum_t w_t x_t & \sum_t x_t^2 \end{pmatrix}^{-1}$$

となるので，完全な多重共線性の場合は，推定値の分散が無限大となる。推定値の分散が無限大という意味は，どこに推定値があるか分からないということを意味する。

14　不均一分散 (不等分散)

回帰式が

$$y_t = \alpha + \beta x_t + u_t$$

の場合を考える。x_t が外生変数，y_t は内生変数，u_t は互いに独立な同一の分布を持つ攪乱項 (最小二乗法に必要な仮定) とする。「独立な同一の分布」の意味は「攪乱項 u_1, u_2, $\cdots$, u_T はそれぞれ独立に平均ゼロ，分散 σ^2 の分布する」である。

分散が時点に依存する場合，代表的には，分散が他の変数 (例えば，z_t) に依存する場合，すなわち，u_t の平均はゼロ，分散は $\sigma_*^2 z_t^2$ の場合は，最小二乗法の仮定に反する。そのため，単純には，$y_t = \alpha + \beta x_t + u_t$ に最小二乗法を適用できない。以下のような修正が必要となる。

$$\frac{y_t}{z_t} = \alpha \frac{1}{z_t} + \beta \frac{x_t}{z_t} + \frac{u_t}{z_t} = \alpha \frac{1}{z_t} + \beta \frac{x_t}{z_t} + u_t^*$$

このとき，新たな攪乱項 u_t^* は平均ゼロ，分散 σ_*^2 の分布となる (すなわち,「同一の」分布)。

$$\mathrm{E}(u_t^*) = \mathrm{E}\left(\frac{u_t}{z_t}\right) = \left(\frac{1}{z_t}\right) \mathrm{E}(u_t) = 0$$

u_t の仮定 $\mathrm{E}(u_t) = 0$ が使われている。

$$\mathrm{Var}(u_t^*) = \mathrm{Var}\left(\frac{u_t}{z_t}\right) = \left(\frac{1}{z_t}\right)^2 \mathrm{Var}(u_t) = \sigma_*^2$$

u_t の仮定 $\mathrm{Var}(u_t) = \sigma_*^2 z_t^2$ が最後に使われている。

よって，$\dfrac{y_t}{z_t}$, $\dfrac{1}{z_t}$, $\dfrac{x_t}{z_t}$ を新たな変数として，最小二乗法を適用することができる。

不均一分散の検定について

$$\widehat{u}_t^2 = \gamma z_t + \epsilon_t$$

を推定し，γ の推定値 $\widehat{\gamma}$ の有意性の検定を行う (通常の t 検定)。

z_t は回帰式に含まれる変数でもよい。例えば，u_t の平均はゼロ，分散は $\sigma_*^2 x_t^2$ の場合，各変数を x_t で割って，

$$\frac{y_t}{x_t} = \alpha \frac{1}{x_t} + \beta + \frac{u_t}{x_t} = \alpha \frac{1}{x_t} + \beta + u_t^*$$

を推定すればよい。β は定数項として推定されるが，意味は限界係数 (すなわち，傾き) と同じなので注意すること。

15　関数型について

$$y_t = \alpha + \beta x_t + u_t$$

この場合，

$$\beta = \frac{\mathrm{d}y_t}{\mathrm{d}x_t}$$

なので，β は，x_t が一単位上昇 (下落) したとき，y_t は何単位上昇 (下落) するのかを表す。すなわち，β は限界係数と呼ばれる。

$$\log(y_t) = \alpha + \beta \log(x_t) + u_t$$

この場合，

$$\beta = \frac{\mathrm{d}\log(y_t)}{\mathrm{d}\log(x_t)} = \frac{\dfrac{\mathrm{d}y_t}{y_t}}{\dfrac{\mathrm{d}x_t}{x_t}} = \frac{100\dfrac{\mathrm{d}y_t}{y_t}}{100\dfrac{\mathrm{d}x_t}{x_t}}$$

となる。

2 つ目の等号では，$\dfrac{\mathrm{d}\log(y_t)}{\mathrm{d}y_t}=\dfrac{1}{y_t}$ が利用される。

3 つ目の等号の分子 $100\dfrac{\mathrm{d}y_t}{y_t}$ や分母 $100\dfrac{\mathrm{d}x_t}{x_t}$ は上昇率を表す。

したがって，β は，x_t が 1%上昇 (下落) したとき，y_t は何%上昇 (下落) するのかを表す。β は弾力性と呼ばれる。

16　同時方程式 (連立方程式) について

単純なマクロ経済・モデル

国内総支出の定義式：

$$GDE_t = C_t + I_t + G_t + E_t - M_t$$

消費関数：

$$C_t = \alpha + \beta Yd_t + \gamma C_{t-1}$$

国民所得の定義式：

$$Y_t = Yd_t + T_t$$

国内総生産の定義式：

$$GDP_t = Y_t + D_t + \epsilon_t$$

三面等価の原則：

$$GDP_t = GDE_t$$

ただし，記号は以下の通り。

GDE_t:	国内総支出
C_t:	家計最終消費支出
I_t:	国内総資本形成 (投資)
G_t:	政府最終消費支出
E_t:	財貨・サービスの輸出
M_t:	財貨・サービスの輸入
Yd_t:	家計国民可処分所得
Y_t:	国民所得
T_t:	直接税
GDP_t:	国内総生産
D_t:	固定資本減耗
ϵ_t:	統計上の不突合

いくつか変数を消去して，3 本のマクロ・モデルを作る。

国内総支出の定義式：

$$GDE_t = C_t + (I_t + G_t + E_t - M_t)$$

消費関数：

$$C_t = \alpha + \beta Yd_t + \gamma C_{t-1}$$

可処分所得の定義式：

$$Yd_t = GDE_t - (T_t + D_t + \epsilon_t)$$

モデルの妥当性のチェック (部分テスト，最終テスト)

1. 個々の推定式のチェック (ここでは，消費関数の推定結果のチェック) $\Longrightarrow$ 部分テスト

2. 連立方程式としてのマクロ・モデルの妥当性のチェック $\Longrightarrow$ 最終テスト

シミュレーション

1. 政府支出増加の国内総支出への影響 (短気的効果と長期的効果)

2. 直接税率増加の国内総支出への影響 (短気的効果と長期的効果)