時系列を単純法で解く

Masato Shimura
jcd02773@nifty.ne.jp

2005 年 6 月 24 日

目次

1 AR 1

2 VAR 2
2.1 経過と説明 3
2.2 TIMSAC のデータで 5
2.3 AIC 7

1 AR

時系列のアルゴリズムにはユール · ウオーカー法、バーク法があり、最小二乗法の解法として、ハウスホルダー法が用いられている。

ここでは鈴木「J 言語による統計解析」(1996) にならってハウスホルダー変換を用いない直説法で解く。

3 次の組み合わせの例。時系列データを次のように組み合わせる。Y は対応しない次数分を上から落とす。X は最終行が対応しないので落とす。これは予測に用いる。

Y	X		
[1]	(0)	(0)	(0)
[2]	(1)	(0)	(0)
[3]	(2)	(1)	(0)
4	3	2	1
5	4	3	2
6	5	4	3
7	6	5	4
8	7	6	5
9	8	7	6
10	9	8	7
11	10	9	8
12	11	10	9
(X)	[12]	[11]	[10]

2 VAR

多変量時系列はユール・ウオーカー法がネスティッドアレーになり、レビンソンのアルゴリズムを用いて解くことになる。

これを一変数と同じようにハウスホルダー変換を経ないで直接解く方法を考える。
多変数を指定次数に従い次の様な組み合わせにして、時間の応答を合わせる。ここでは 3 次の例とした (次数毎に変数をすらすと過剰に回帰する。)

Y を次数分頭からカットして、X の組み合わせた最後行は、対応しないので、予測に回す。

ここで $\dfrac{Y}{X}$ とすれば多変量時系列の回帰係数が求まる。

Y			X								
11	21	31	X_1			X_2			X_3		
12	22	32	t_{-1}	t_{-2}	t_{-3}	t_{-1}	t_{-2}	t_{-3}	t_{-1}	t_{-2}	t_{-3}
13	23	33	13	12	11	23	22	21	33	32	31
14	24	34	14	13	12	24	23	22	34	33	32
15	25	35	15	14	13	25	24	23	35	34	33
16	26	36	16	15	14	26	25	24	36	35	34
17	27	37	17	16	15	27	26	25	37	36	35
18	28	38	18	17	16	28	27	26	38	37	36
19	29	39	19	18	17	29	28	27	39	38	37
20	30	40	20	19	18	30	29	28	40	39	38
21	31	41	21	20	19	31	30	29	41	40	39
22	32	42	22	21	20	32	31	30	42	41	40

一行で書けば次のようになる。

```
var_simple2=:4 : '(x.}.tmp) %. L:0 |."1 L:0 }: L:0 x.>\ L:0 {|:tmp=:dev2 y.'
```

2.1 経過と説明

```
   11 21 31 +/ i. 11   NB. transpose is |:
11 12 13 14 15 16 17 18 19 20 21
21 22 23 24 25 26 27 28 29 30 31
31 32 33 34 35 36 37 38 39 40 41
```

```
var_sub0=:4 : 0
|."1 L:0 x.>\ L:0 {|: y.
)
```

変数毎に box に入れて general array にして、box 内を 3 個までずらす。

```
3>\ L:0 {( 11 21 31 +/ i. 11)
+--------+--------+--------+
|11 12 13|21 22 23|31 32 33|
|12 13 14|22 23 24|32 33 34|
|13 14 15|23 24 25|33 34 35|
|14 15 16|24 25 26|34 35 36|
|15 16 17|25 26 27|35 36 37|
|16 17 18|26 27 28|36 37 38|
|17 18 19|27 28 29|37 38 39|
|18 19 20|28 29 30|38 39 40|
|19 20 21|29 30 31|39 40 41|
+--------+--------+--------+
```

ボックス内を回転させ、t_{-1}, t_{-2}, t_{-3}) となるようにする。

```
3 var_sub0 (|: 11 21 31 +/ i. 11)
+--------+--------+--------+
|13 12 11|23 22 21|33 32 31|
|14 13 12|24 23 22|34 33 32|
|15 14 13|25 24 23|35 34 33|
|16 15 14|26 25 24|36 35 34|
|17 16 15|27 26 25|37 36 35|
|18 17 16|28 27 26|38 37 36|
|19 18 17|29 28 27|39 38 37|
|20 19 18|30 29 28|40 39 38|
|21 20 19|31 30 29|41 40 39|
+--------+--------+--------+
```

このサンプルは行列式の値が 0 になり線形従属なので、逆行列が求められない。

-/ . *　　行列式の値を求める

2.2　TIMSAC のデータで

統計数理研究所の提供する FORTRAN の時系列プログラム TIMSAC にサンプルデータが付いている。

AIC はまだユールウオーカー法の借り物だが、12 次を最適としている。

(長いので、3 次を例示する。)

```
7.3 ": (L:0)  3 var_simple data
```

4 次の場合である。BOX はスペースの関係で縦に重ねた。上から X_1, X_2, X_3 である。

BOX の中は、縦は順に $t_{-1}, t_{-2}, t_{-3} t_4$ となっており、横は X 相互関係である。

```
,.  4 var_simple data
+--------------------------------+
|   1.73511  0.355886 0.0765287  |
| _0.837703 _0.251929  0.858603  |
|  0.116042 _0.699268   2.30439  |
|_0.0345968  0.960708  _2.23717  |
+--------------------------------+
|  0.447935    1.16684  _1.62722 |
|_0.0849224 _0.0485796   2.30718 |
|_0.0523771  0.0028819  0.342791 |
| _0.129053  _0.133072 _0.967287 |
+--------------------------------+
|   0.201807  0.0150937    1.0604|
|_0.00801159 0.00747033 _0.519573|
|  0.0148035 _0.0140778 _0.211859|
|   0.178705   0.050196  0.291256|
+--------------------------------+
```

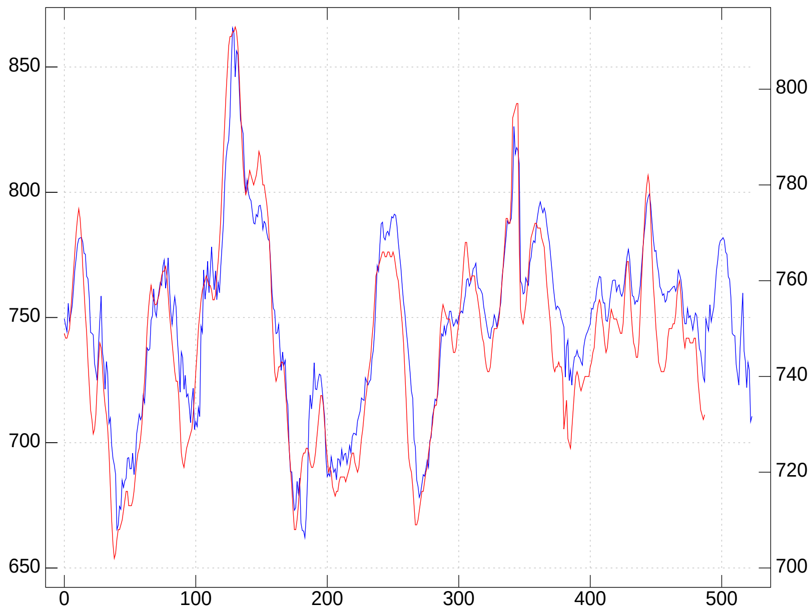

図 1　data0

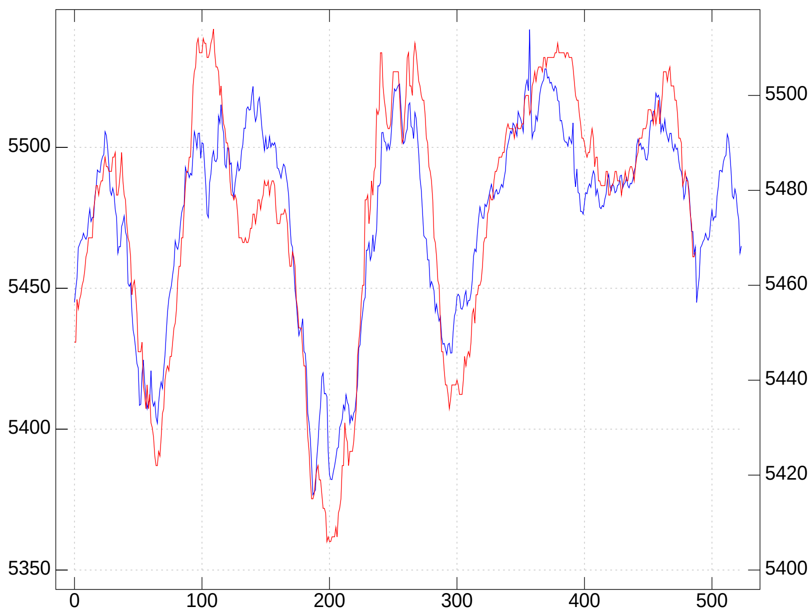

図 2　data1

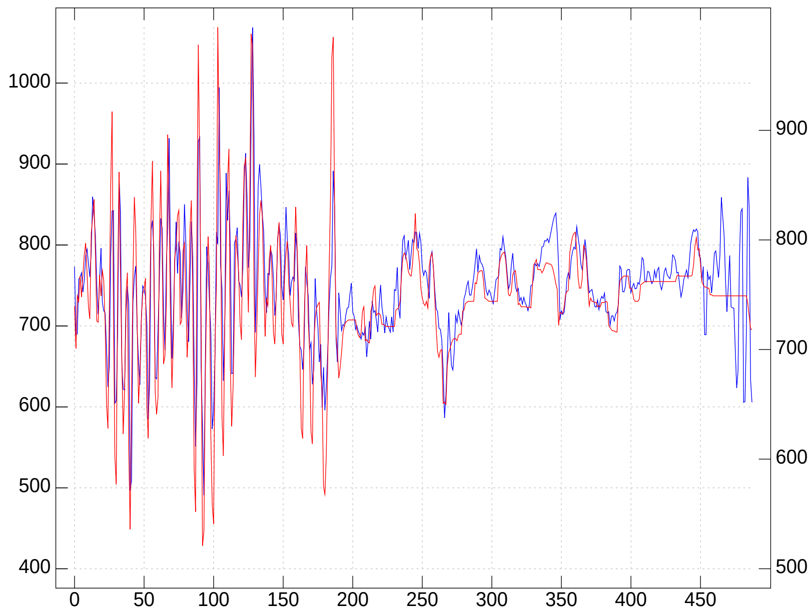

図 3　data2

2.3　AIC

$AIC(M) = -2 \times MLL(M) + 2 \times k$

重相関モデルは

$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k$

残差平方和は

$Q = \sum_{n=1}^{n} \{y_i - (b_0 + b_1 x_1 + \cdots + b_k x_{ki}\}^2$

$= (y - Xb)\,'(y - Xb)$ を最小にする。

ここで Q を最小にする b は

$b' = (y'X)(X'X)^{-1}$

残差平方和は

$Q = (y - Xb)'(y - Xb) = y'y - (y'X)b$

$MLL = -\frac{1}{2}\log\frac{Q}{n}$

$AIC = n\log\frac{Q}{n} + 2 \times (k+1)$ である。

多変量自己回帰では b は次数分のマトリクスになり、Q が次数分できる。

ハウスホルダーの場合の AIC は

$\hat{\sigma}_j^2(1) = \dfrac{1}{N-m} \sum_{i=kj+1}^{km+1} s_{i,km+1}^2$
$AIC_j(1) = (N-m)(\log 2\pi\hat{\sigma}_j^2(1)+1) + 2(kj+1)$
$\hat{\sigma}_j^2(k) = \dfrac{1}{N-m} \sum_{i=kj+k}^{km+k} s_{i,km+k}^2$
$AIC_j(k) = (N-m)(\log 2\pi\hat{\sigma}_j^2(k)+1) + 2(kj+k)$
として次数分求めている。
AR の MLL は $MLL = -\dfrac{n-k}{2}(\log Q_k/(n-k))k$
$AIC = (n-k)\ times MLL + 2 \times k$
$= (n-k) \times (\log Q_k/(n-k)) + 2 \times k$
で求められる。