

## 第十二章 MicroTSP

### §12.1 MicroTSP 入门

#### §12.1.1 简介

MicroTSP 是基于大型机TSP软件包的微机产品。TSP的开发始于五十年代，主要用于时间序列分析，现已成为计量经济学和时序分析主要的软件包。TSP也有相应的微机版本。MicroTSP 于1981年在Apple II 计算机上引入，1982年推出第一版，Micro TSP 6.5 版1989年开始投放软件市场，能力更强，使用更方便。

MicroTSP 提供了计量经济学和预测工作中实用的统计技术，其应用领域包括财务分析、经济研究、销售预测、费用分析、宏观经济预测、利率预测以及科学数据分析和评价。MicroTSP 方便易学、操作简便。

统计工具：描述统计、多元回归——普通最小二乘、加权最小二乘、多项式分布滞后(PDL)、ARIMA - Box-Jenkins (包括季节项)、ARMAX (带ARMA误差的结构估计)、相关图、非线性最小二乘、两阶段最小二乘、三阶段最小二乘——线性与非线性、似乎不相关回归(Seemingly Unrelated Regression) - 线性与非线性、Probit 和Logit 模型、向量自回归、VAR 冲激响应和方差分解、季节调整。

预测工具：预测、预测方程库、系统估计和VAR 模型求解和模拟(线性和非线性)、指数平滑、随机模拟方法。

数据管理：数据编辑、访问其它标准格式文件、时间转换等。

高分辨率图形：时序数据图示、散点图、条图、圆图、直方图、双标度图示、对数尺度图标、冲激响应函数。图形的软硬输出等。

其它功能：表格和统计输出可以打印或以文件形式存贮、批处理程序和宏调用、编辑、显示和管理TSP 模型和程序文件、支持数学协处理器。

MicroTSP 大部分功能由菜单驱动或MicroTSP行命令完成，如：

**DATA** 一用于输入、扩充和修改时间序列数据。

**GENR** 一在已有序列的基础上，利用各类型的计算公式，生成一个新的序列。

**PLOT** 一在屏幕或打印机上生成高分辨率的趋势图。

**SCAT** 一生成高分辨率的散点图(相关图)。

**BAR** 一生成高分辨的直方图。

**PIE** 一生成高分辨率的饼图。

**PRINT** 一在屏幕上显示序列。

**LS** 一最小二乘法(多元回归)。

**AR** 一带有自回归误差校正项的最小二乘法。

**TSLs** 一两阶段最小二乘。

**NLS** 一非线性回归。

**PROBIT, LOGIT** 一二分类数据建模。

**SYS** 一估计线性或非线性方程组。

**VAREST, VARSTAT** —估计和分析向量自回归方程组。

**COVA, IDENT, CROSS, HIST** —描述统计量: 相关系数, 协方差, 自相关系数, 互相关系数以及在识别时间序列模型等方面非常有用的频率直方图。

**AR, MA** —在方程的估计中包含自回归和移动平均。

**SAR, SMA** —在季节变动分析中包含自回归和移动平均。

**PDL** —多项式分布滞后处理。

**FORCST, FIT** —利用已有的回归方程计算预测值。

**SOLVE** —求解联立模型并计算模拟值。

**CREATE** —建立内存工作文件。

**SAVE** —存储工作文件到磁盘。

**LOAD** —把磁盘上的工作文件加载到内存。

**READ, WRITE** —读写Lotus 格式(.WKS, WK1, .WR1.等) 文件, Lotus 打印格式(PRN) 文件, DIF 格式文件和其他格式的文件。

**EDIT** —建立和修改模型和其他文本文件。

**F2** —命令回顾。

MicroTSP 软件提供了五个样本供学习使用。

### §12.1.2 运行

MicroTSP 有四种工作方式:

- 菜单方式。利用系统提供的四个菜单, 通过功能键F3—F6完成操作;
- 行命令方式。每输入一个命令, 系统询问该命令其它信息, 引导后续操作;
- 行命令带参数输入方式。用户把操作命令及参数一次输入, 提交系统完成;
- 批处理方式。用户使用MicroTSP 本身的编辑程序或其他的字处理程序建立MicroTSP 的命令文件, 供MicroTSP运行。

第一种是新用户常常使用。当对系统比较熟悉后, 便可以使用第二、三种方式。批处理方式主要是便于用户把常规的处理过程变成一个批处理作业, 反复使用。在批处理作业运行期间, 用户可以做其他事情, 而不必始终坐在显示器前。

使用命令TSP 启用系统后, 出现系统提示:

>

用户既可以输入一个单独的命令, 也可以输入包含命令及其选择参数的命令参数表, 还可以通过功能键F3-F6 进入菜单操作。如采用菜单操作, 可根据屏幕下方的提示:

F1—中断 F2—命令重现 F3—文件 F4—数据 F5—统计 F6—TSP 控制

F3—F6 功能键用于选择TSP 软件包的四个主要菜单。按F3 选择文件菜单,

屏幕上的菜单格式为:

## 文件操作

- (1) 工作文件 (开始会话)
- (2) 数据库文件操作
- (3) 显示磁盘目录            DIR
- (4) 改变子目录            CD
- (5) 编辑文本文件        EDIT
- (6) 文件改名            REN
- (7) 删除文件内容        DEL
- (8) 显示文件            TYPE
- (9) 读外部文件        READ
- (A) 写外部文件        WRITE

## F1 中断(F3-F6 选菜单)

开始时选择菜单的第一行。选择可以有两种方式，其一是输入与有关命令对应的数字或字母；其二是利用键盘右面的光标移动键将光标移动到有关命令行上，然后按回车键。现输入1，表示开始会话过程。

此时将出现工作文件操作菜单：

- (1) 在内存中建立工作文件    CREATE
- (2) 加载工作文件到磁盘    LOAD
- (3) 存储工作文件到磁盘    SAVE
- (4) 扩充工作文件样本区间    EXPAND
- (5) 按序列名排序        SORT

## F1 中断(F3-F6 选菜单)

因为要建立工作文件，在第一行按回车，然后屏幕上会出现数据频度提示：

数据频度

- (U) 非时间数据
- (A) 年度数据
- (Q) 季度数据
- (M) 月度数据

## F1 中断—中止程序运行

在这里输入A，表示在工作文件中输入年度数据。

开始时间? //1970

结束时间? //1980

// 后是回答的内容，这时屏幕上方的状况窗口的有关数据已被更新了，而且系统停止了提问并给出命令提示符“\_”表示用户可以输入其他的命令了。

现在可以利用数据编辑功能输入数据了，按F4 选择数据管理菜单，则数据管理的菜单内容会显示在屏幕上：

## 数据管理

- |                 |       |
|-----------------|-------|
| (1) 调整样本区间      | SMPL  |
| (2) 利用已有序列生成新序列 | GENR  |
| (3) 数据编辑        | DATA  |
| (4) 季节调整        | SEAS  |
| (5) 序列编组操作      | GROUP |
| (6) 内存中序列的改名    | R     |
| (7) 内存中序列的删除    | D     |
| (8) 图形操作        |       |
| (9) 显示数据报表      | SHOW  |
| (A) 打印数据报表      | PRINT |

F1 中断(F3—F6 选菜单)

输入3 选择数据编辑。TSP软件包的数据编辑程序开始运行。当软件包提示用户输入序列名时，即可输入您要输入的变量名。设为REV 告诉数据编辑程序为输入序列REV 做好准备。于是屏幕上显示初始日期为1970，并且在下边显示：

obs REV

光标定位在1970 右边一个高亮度区域内，提示用户在此输入样本观察值。每当用户按一次回车键，数据编辑程序都修改一个样本观察值并做好接受另一个观察值的准备。

假定用户在REV 序列中输入了如下观察值：

```
1970  697
1971  814
1972  963
1973 1122
1974 1224
1975 1369
1976 1539
1977 1780
1978 2161
1979 2605
1980 2915
1981  X
```

X 表示退出数据编辑系统，此时出现命令提示符。

若输入数据有误，在数据编辑中可以用命令进行修改，屏幕上方有关于这些命令的解释说明。例如，按回车键后，发现刚刚输入的观察值错误，可以输入命令B 使光标返回到上一个观察点，以修改刚刚输入的观察值。也可以用命令Ni 使光标定位到由 i 表示的观察点上，例如输入命令N75 后，光标定位到序列中1975 年的观察值上，如果要编辑1980 年的观察值，可以输入命令：N80。

数据编辑操作或其他操作之后，TSP 都会在屏幕上方显示最新的状态。现在序列目录中已有了REV。假如在REV 序列输入后发现序列中1976 年的观察值错误，可以调用数据编辑功能进行修改，这次不使用菜单方式，而直接输入DATA 命令，当软件包询问序列名时，输入：REV

按回车键，屏幕上将显示REV序列的观察值，输入命令：N76后，光标定位在1976年的观察值上。键入正确的观察值后按回车键，序列中1976年的观察值被新的观察值代替。输入X字符后结束对REV序列的编辑。

下面再输入一个变量GNP，这次将数据编辑命令与序列名一起输入，即输入：

DATA GNP

其区间与REV完全相同，顺序输入下列数据值：

992.7,1077.6,1185.9,1326.4,1434.2,1549.2,1718.0,1918.0,2156.1,2413.9, 2627.4

最后键入X结束。TSP软件包允许用户利用已有序列的方程式生成新的序列。按F4调出数据管理菜单，然后选择2，即GENR命令，则TSP提示：

方程式为？//RATIO=REV/GNP

这时在内存中生成了RATIO序列，并且序列名出现在状态窗口中。

GENR命令用于转换数据，用户可以说明十分复杂的代数公式，计算结果构成由方程中等号左边的名字命名的序列，如果由该序列命名的序列已经存在，新生成的观察值将代替序列中原有的观察值，否则将建立一个新的序列。

GENR命令也可以用于计算标量和统计函数。

现在我们已经建立了一个新序列RATIO。为了显示它请按F4功能键调出数据管理菜单，然后选择SHOW命令。当系统询问序列名时回答RATIO，序列中的数据显示在屏幕上的表格区内：

obs	RATIO
1970	0.702125
1971	0.755382
1972	0.812041
1973	0.845899
1974	0.853437
1975	0.883682
1976	0.895809
1977	0.928050
1978	1.002273
1979	1.079167
1980	1.109462

按任意键后返回初始屏幕，处于系统提示下。

新序列生成后，可以进行各种处理如打印输出、屏幕显示，制图，回归分析、预测或存入磁盘，等等。

现在按F4并选择SMPL命令，用

1971 1980

回答系统提问的样本区间，则SHOW等操作就不使用1970年的观察值了。

利用TSP软件包可以非常方便地研究滞后序列。例如在命令提示符后直接输入GENR命令和方程：

GENR PCHR=100\*(RATIO-RATIO(-1))/RATIO(-1)

这里RATIO是假日旅馆收入REV与国民生产总值GNP的比率，RATIO(-1)是RATIO的滞后序列，滞后期为1。由上面方程计算的PCHR为RATIO的增长百分比。输入SHOW命令可以显示新生成的序列PCHER。

按F6 选择TSP 控制菜单，则显示出下列命令菜单：

TSP 控制命令

- |                 |        |
|-----------------|--------|
| (1) 结束会话/退回DOS  | EXIT   |
| (2) 运行TSP批处理程序  | RUN    |
| (3) 运行DOS 命令    | SYSTEM |
| (4) 打印机状态参数设置   |        |
| (5) 装配TSP运行环境   | CONFIG |
| (6) 参数选择及对应值    | OPTION |
| (7) 内存使用报告      | FREMEM |
| (8) 清屏并显示序列目录信息 |        |

F1 中断(F3—F6 选菜单)

选择1，则TSP 提问：

放弃当前内存工作文件吗？(Y/N) //Y

则退出TSP 软件包，返回到DOS 命令状态。也可用行命令EXIT 退出。使用SYSTEM命令可以暂时退回到DOS系统，DOS操作结束后用EXIT返回MicroTSP。

## §12.2 MicroTSP 数据分析

### §12.2.1 显示时间序列图形

如果用户的计算机配有图形设备，则可以用PLOT, SCAT, BAR 和PIE 命令生成高分辨率的图形。如果用户的硬件不能生成图形，则不能使用这些命令。

按F4 调用数据管理菜单。选择第8 项功能，则显示图形命令菜单：

- |              |         |
|--------------|---------|
| (1) 曲线图（趋势图） | PLOT    |
| (2) 相关图      | SCAT    |
| (3) 直方图      | BAR     |
| (4) 圆图       | PIE     |
| (5) 频度直方图    | HIST(G) |
| (6) 加载图形文件   | LGRAPH  |
| (7) 打印图形文件   | PGRAPH  |

F1 中断(F3—F6 选菜单)

PLOT 命令显示时间序列图。横轴是时间方向，各个数据点之间以直线相连。选择PLOT后，出现比例尺参数设定方法菜单：

- (A) 自动设定—单比例尺
- (M) 手工设定—单比例尺
- (N) 标准化图
- (D) 双比例尺—图不交叉
- (X) 双比例尺—图交叉
- (R) 标准差标出的残差图
- (S) 参数设定选择

F1 中断—中止程序运行

输入A 自动设定比例尺。TSP 询问序列名，这时输入：

RATIO

然后会在屏幕上看到高分辨率的曲线图。在观看完图后，输入X，返回命令接受状态。SCAT 命令生成两个序列间的相关图。这次不使用菜单驱动，直接输入命令SCAT，TSP

提问

序列名表? // REV GNP

系统会给出图形选择:

散点图选择

(S) 简单散点图

(C) 邻近两点连接式

(R) 给出回归线

(B) 邻近两点连线并给出回归线

F1 中断—中止程序运行

选择S，TSP 将以GNP 为X 轴方向，REV 为Y 轴方向给出两个序列的散点相关图。在看完图后，输入X，返回系统命令接受状态。

TSP 能够进行普通最小二乘法、两阶段最小二乘法、带有自回归误差校正的最小二乘法估计。在所有估计中，用于估计的样本点都由当前的SMPL 确定。

### §12.2.2 统计量的计算

COVA 命令计算并打印多个序列各自的均值、最大、最小值、标准偏差及它们的协方差矩阵和相关矩阵。例如可以输入命令:

COVA CONS GNP G R GNP(-1)

如果给出W选择，COVA将在计算描述统计量之前对每个序列进行加权处理。命令输入方式为:

COVA(W=POP) CONS GNP G R GNP(-1)

若使用了M 选择，COVA只显示均值，标准差和最大值、最小值。例如:

COVA(M) CONS GNP G R GNP(-1)

IDENT命令计算一个序列的自相关系数和偏相关系数。输入IDENT命令后，软件包要求给出序列名和要计算的相关系数的数目。另一种命令输入方式是在IDENT后面的括号中给出要计算的相关系数的数目，然后给出序列名。例如输入命令:

IDENT

IDENT(12) IBMPRICE

IDENT 计算出自回归系数的标准差。还计算Box-Pierce 的Q-检验值，即自相关系数的平方和。Q-检验值可用于进行自相关系数均为零的假设检验，即假定序列为白噪声。在零假设检验下，如果事先没有对序列进行ARIMA 分析，Q- 统计量服从 $\chi^2$ -分布，自由度为自相关系数的数目。如果序列是检验ARIMA 估计的残差，则自由度为自相关系数的个数减去事先估计时的自回归和移动平均的项数。欲知详细的信息，请参阅Box 和Pierce (1970) [1]。

CROSS 命令计算两个序列的相关系数。输入CROSS 命令后，软件包提示给出要计算的相关系数的数目和序列名。也可以像IDENT 命令一样，在命令名后的括号中给出要计算的相关系数的数目，然后给出序列名。例如输入命令:

CROSS

CROSS(8) TB3 ASQ

HIST命令计算一个序列的频度图。在频度图中显示数据的频率分布。它将序列的数值范围(由最小值和最大值确定的数值区间) 等分成一些小区间，然后显示落入每个小区间的

观察值数目。作为隐含值，TSP 将数值区间等分成10 个左右小区间。但是通过在HIST 后面的括号中说明小区间的长度能改变系数的隐含区间分割方法。例如：

```
HIST RESID
HIST(2.5) AGE
```

标准的频度图输出是用文本字符格式，并同时显示序列的均值、标准差、最大最小值。用户可以输入 P 打印输出结果。若有图形显示器，可以输入 G 显示高分辨率的频度直方图。

包括 G 选择参数的 HIST 命令会立即在屏幕上以图形方式显示频数图。G 选择也可以和区间规格参数一起使用。若没有指出区间规格，TSP 大致划分 30 个左右的小区间来显示图形。输入例子如下：

```
HIST(G)
HIST(G,.05) IQ
```

### §12.2.3 回归分析

按 F5 选择统计运算菜单，则具体的内容显示如下：

统计运算和模型模拟

- |                 |        |
|-----------------|--------|
| (1) 描述统计和统计检验   |        |
| (2) 单方程估计方法     |        |
| (3) 估计方程处理和预测分析 |        |
| (4) 系统（文件）估计    | SYS    |
| (5) 向量自回归       | VAR    |
| (6) 求解模型（文件）    | SOLVE  |
| (7) 编辑系统文件或模型文件 | EDIT   |
| (8) 指数平滑        | SMOOTH |

F1 中断(F3-F6 选择菜单)

选择 2，则显示出单方程估计方法菜单：

单方程估计方法

- |                         |           |
|-------------------------|-----------|
| (1) 可带有 ARMA 项的普通最小二乘法  | LS        |
| (2) 带有异方差性处理的普通最小二乘法    | LS (H)    |
| (3) 可带有 ARMA 项的两阶段最小二乘法 | TLSLS     |
| (4) 非线性回归               | NLS       |
| (5) 加权最小二乘法             | LS (W)    |
| (6) 加权两阶段最小二乘法          | TLSLS (W) |
| (7) 加权非线性最小二乘法          | NLS (W)   |
| (8) 非线性回归 (NLS) 的初值设定   | PARAM     |
| (9) 广义 logistic 模型      | LOGIT     |
| (A) 二元概率分布模型            | PROBIT    |

F1 中断(F3-F6 选择菜单)

现在选择 1，即进行最小二乘估计。系统将开始下面的会话：

请输入被解释变量名？//REV

系统提示：

解释变量命表中可以包含 AR, SAR, MA, SMA 和 PDL 参数项，

请输入解示变量名表？//C GNP

屏幕上显示统计结果。然后TSP提示:

要显示协方差矩阵? (P, S, ENTER)

若回答S, 则会在屏幕上显示协方差矩阵。之后提问

显示残差值, 实际值和拟合值吗? (P, S, ENTER)

回答S, 则显示出残差图(残差图为字符方式, 不需要图形设备), 如果装配了图形设备, 输入G, 则会得到高分辨率的带有残差、实际值和拟合值的残差图。

最后, 系统提问:

重新输出计算结果? (P, S, ENTER)

按回车键, 将结束回归过程, 回到命令提示。

### (一)普通最小二乘法

计算普通最小二乘法, 可使用不进行任何误差处理的LS命令。键入LS命令及一组序列名, 第一个序列是被解释变量, 其余的序列都是解释变量。可以利用普通最小二乘法建立一个简单的消费函数, 即输入命令:

```
LS CONS C GNP
```

TSP 要求回归方程象处理其它解释变量一样处理常数项, 这与许多回归程序是不一样的。在TSP软件包中, 有一个特殊的、预先生成的序列, 名字为C。这个序列可以用于LS及其他统计运算中, 但不必象输入其他数据那样输入该序列, 所以用户不要把自己建立的序列命名为C。

滞后序列可以出现在LS或其他统计运算中。一个序列的滞后序列与原序列同名, 只是在序列名后的括号中给出滞后期。如: LS CONS C CONS(-1) GNP

设Y是被解释变量矩阵, X是解释变量矩阵。LS命令计算最小二乘法回归系数和各种有关的统计量。这些统计量包括系数的标准差和相应的t-检验值、回归标准差、D-W检验值。如果回归方程中包含常数项, 则判定系数( $R^2$ )、调整后的判定系数以及总体F-检验值也将显示出来。TSP提供了几组可选择的参数, 使用户可以灵活地掌握那些回归结果在屏幕上显示或打印。基本回归结果显示之后, 软件包将询问是否查看系数的协方差矩阵。可以输入S在屏幕上显示这个矩阵, 或输入P将它打在打印机上打印出来, 或按回车键跳过这个选择。之后系统提问:

要显示残差值、实际值和拟合值吗? (P,S,G,ENTER)

回答P或S将打印或显示字符形式的残差图、被解释变量的实际值、拟合值和残差值。回答G将以双比例尺方式生成残差、实际值和拟合值的高分辨率图形, 也可按回车键跳过这个选择。

在LS命令中可以使用两个选择来处理异方差性。H选择要求TSP使用相合的协方差矩阵代替前面给出的计算标准差和t-检验值时使用的协方差矩阵, 请参阅[2]。如: LS(H) EDUC C AGE INCOME FATHER

W选择加上一个等号一个权值序列名说明进行加权最小二乘法。若使用了这个选择, 则在进行估计之前所有数据都要与权值序列相乘。如:

```
LS(W=SCALE) EDUC C AGE INCOME FATHER
```

### (二)两阶段最小二乘法

TSL命令计算两阶段最小二乘模型。两阶段最小二乘法有时也叫工具变量法。在TSL命令中, 被解释变量和解释变量按LS命令中变量的排列方式列出后键入一个@字符, 然后列

出工具变量。例如,在消费方程中,可以把政府支出G、货币供给量的自然对数LM、趋势变量TIME 作为外生变量,因此可用做工具变量,输入命令如下:

```
TSLS CONS C GNP @ C G TIME LM
```

常量C 总是一个合适的工具变量,设Z是工具变量值构成的矩阵,Y和X分别是被解释变量和解释变量矩阵,TSLS根据以下方程计算估计方程的系数:

$$(X'Z(Z'Z)^{-1}Z'X)^{-1}X'X(Z'Z)^{-1}Z'Y$$

这些系数的协方差矩阵为:  $XZ(X'Z(Z'Z)^{-1}Z'X)^{-1}$

与LS命令的计算过程一样,有关统计量也被计算。所有与残差有关的统计量的计算公式与LS命令的计算公式一样。这里的残差是“结构残差”  $Y-Xb$ , 而不是“第二阶段残差”。在TSLS 命令中可以使用W选择来进行加权两阶段最小二乘法。

注意:进行TSLS 估计时,即使方程中有常数项,  $R^2$ 也可能是负值。

### (三)一阶序列相关

当线性回归模型的扰动项序列相关时,由普通最小二乘法估计出的系数尽管是无偏的,却是无效的。当扰动项出现一阶序列相关时,AR(1)命令提供了得到有效估计值的方法,为了使用序列相关校正方法估计消费函数,输入命令:

```
LS CONS C GNP AR(1)
```

用AR(1)方法进行误差序列校正的LS 计算过程使用了D.Cochrane 和G.H. Orcutt [4] 提出的两阶段迭代方法。这种方法利用普通最小二乘法计算出的残差估计rho 的值,对被解释变量进行变换,使得变换后的估计方程的残差项基本上序列不相关,然后使用变换后的变量估计回归方程。变换方程是:

$$xi=xt-rho*xt-1$$

这种计算过程一直重复进行到rho收敛或迭代次数达到了预先给定的上限为止。一般来说, Cochrane-Orcutt 过程渐近于极大似然法,但用小样本估计时它们是不同的。

包含AR(1) 命令的LS命令也给出通常的回归计算结果和根据rho 变换变量计算各种统计量及残差散点图。如果一阶AR 校正误差序列是恰当的,输出的残差序列将是不相关的白噪声序列。

对于包含AR(1)命令的LS命令,用SMPL 命令确定的样本区间一定要保证Cochrane-Orcutt过程使用的滞后观察值是存在的。例如,如果被解释变量和解释变量的第一个观察值对应的观测点均为1949 年,则SMPL 命令说明的样本期最早只能从1950年开始。如果以工作文件的第一个样本点为起点,包含AR(1)命令的LS 命令会出错。

使用OPTION 命令,可以在执行LS 命令时控制迭代过程。通常,迭代20 次后,即使rho不收敛,也将停止执行LS 命令。选择MAXIT 能够改变对迭代的次数限制。例如,为将迭代次数限制增为100,只需输入命令:

```
OPTION MAXIT 100
```

LS 通过检查两次迭代得到的rho估计值之差来决定收敛程度。通常,当差值小于等于0.005 时,LS 停止迭代并打印计算结果。为改变收敛判别标准,应选择CONVERGE。例如,为提高计算精确性,把判别标准改为0.001,应输入命令:

```
OPTION CONVERGE 0.001
```

每一次迭代时,LS 试验给出步长。它首先检验采用这个步长能否降低残差平方和,如果残差平方和降低,就采用它并检验收敛程度,如果需要,就进行下一次迭代。如果所取的

步长不能降低残差平方和,取步长的一半执行迭代过程。假如还不能降低残差平方和,再把步长一分为二。如果此时仍无改进,LS 命令停止计算,并给出不能继续计算的信息。

MAXAZQ 选择控制LS 将步长一分为二的次数。通常只分两次,但命令  
OPTION MAXSZQ 4

将其增加到4次。一旦用OPTION 命令改变了一个参数,该参数在本次会话中一直有效。可以用一个OPTION 命令重新定义多个参数,如:

OPTION CONVERGE .01 MAXIT 50

#### (四)多项式分布滞后

估计方程中可以包含多项式分布滞后,一个方程中最多可以含有五个PDL项。每一项通知TSP 对一个序列计算多项式分布滞后系数,并根据多项式的形式确定这些系数。

PDL 项后面的括号中应给出一个序列名,如:LS SHPMNT C PDL(ORDERS) 表示SHPMNT 被ORDERS 序列的分布滞后拟合。输入以上命令后,TSP 要求给出分布滞后的补充信息。必然提供:

滞后长度—即滞后多少时间周期。

多项式的阶数。

是否强制系数在滞后分布的开始或末端趋近于零,或在开始和末端都趋近于零。用以下数字表示对估计系数的限制:

- 1— 在分布近端趋近于零;
- 2— 在分布远端趋近于零;
- 3— 在分布近端和远端都趋近于零。

也可以在括号中序列名后面给出这些信息。如果不做任何限制,可以省略这些信息。例如输入命令:

LS SHPMNT C PDL(ORDERS,8,3)

LS SHPMNT C PDL(ORDERS(-1),12,4,2)

PDL 命令也可以用在TSLs 命令中。如果PDL 后面括号中给出的序列是外生的,应做为工具变量。为此,可以给出PDL(\*),这意味着所有的PDL变量做为工具变量。例如可以输入命令:

TSLs SHPMNT C TBILL MI PDL(ORDERS(-1),12,4,2)

@ FED FED(-1) PDL(\*) 在PDL 的计算过程中,TSP 临时产生一些名为PDL1,PDL2,⋯的序列。这些序列用在LS 或TSLs 估计中。如果对PDL 估计很熟悉,可以将这些序列用作其它目的,例如作为TSLs 中的一个工具变量。

### §12.2.4 高级统计技术

高级统计技术包括非线性最小二乘法和向量自回归模型。

#### (一)非线性最小二乘法

非线性模型有两种:一种是关于变量是非线性的;另一种是关于参数是线性的,当然也可以是关于二者均是非线性的。若一个方程关于变量是非线性的,但关于参数是非线性的,则能够用LS或TSLs方法来估计。例如,若有模型:

$Y=D+A*\text{LOG}(L)+B*\text{LOG}(K)$  关于变量L 和K 是非线性的,但关于参数是线性的,这样可以通过建立新的变量来使用普通最小二乘法:

GENR LOGL=LOG(L)

GENR LOGK=LOG(K)

LS Y C LOGL LOGK

但是这种方法只适用于方程的系数是线性的场合。当方程的系数是非线性的情况时,如方程:

$$Y=D*(L^A)*(K^B)+F*DUM$$

就无法建立新的序列然后用LS去估计。这时我们就可以使用非线性最小二乘法命令NLS。

NLS 使用特殊的方法来表示方程从而使其能够记着哪些参数是要估计的。这些参数依次叫做C(1),C(2),...估计上述方程的NLS 命令为:

$$NLS Y=C(1)*(L^C(2))*(K^C(3))+C(4)*DUM$$

在使用NLS命令前,必须为TSP提供待估计参数的初始值。PARAM 命令就是为此设置的。例如,下面的PARAM 命令就为上述NLS命令提供了一组可用的初值:

PARAM 1 213 2 .7 3 .3 4 0

这些命令设置C(1)=213,C(2)=0.7,...。当NLS 命令估计完成时,估计出的系数还保留在C(1)、C(2)...中。如果这些数值可以用作下次估计的初始值,就不必使用新的PARAM 命令重设初值,如果使用了新的PARAM 命令,仅仅需要说明要修改的参数的初值。

一旦用PARAM 命令设置了参数初值,就可以用NLS 命令进行估计。NLS 估计过程如下:每一次迭代,通过对其中一个参数做少量的改变来计算有关参数的偏差,并查看方程变化了多少。然后将被解释变量与这些偏差做回归。在必要时,为了避免共线性问题,要使用曲线回归,但NLS 的最终结果不包含曲线回归。回归给出一个对参数计划进行修改的数值向量。NLS 评价这些计划修改的参数是否降低了残差平方和。如果残差平方和降低了,就对参数进行修改并开始新一轮迭代。否则就对计划修改量进行分割。一般来说,在进行一定次数的分割之后对方程会有所改善。这种过程一直进行下去,直到修改的数量与参数本身相比非常小为止。

通过修改MAXIT、CONVERGE 和MAXSQZ 的参数值可以控制迭代过程。

有时NLS 甚至在第一次迭代就不能改善残差平方和,这样迭代过程就会停止。在大多数情况中,这种失败原因在于参数的初始值给得不好。例如,在上面的方程中,如果C(1)的初值设为0,则方程与C(2)和C(3)有关的偏差也将是0,这样迭代过程就无法开始了。

另一方面,对于许多行为很好的问题,不管定义了什么样的初值,收敛迅速。因此尝试精选好的初始值是没有必要的。一旦发现了一组能够导致收敛的初始值,就应予以保留。

NLS 在试图对参数修改的操作中一直给用户提示信息。当对计划修改量进行分割而改变了步长时,在屏幕上显示一个\*号。在每一次成功地迭代之后,将显示参数值和残差平方和。

在取得了收敛之后,NLS 将给出与LS 相同的信息。虽然两种方法的理论只是近似的,但所有标准的统计结果和检验值的解释是相同的。

下面的例子是用NLS 估计CES 生产函数的参数:

PARAM 1 1 2 .1 3 .5 4 -1

$$NLS Q=C(1)*(C(3)*L^C(4)+(1-C(3))*K^C(4))^C(2)$$

在NLS 中也可以使用W 或WEIGHT 选择来完成加权非线性回归。

### (二)广义logistic和probit模型

广义logistic和probit模型研究反应为二分类的数据。概率反映一种事件发生的可能性,用0、1之间的一个数来表示。所以说明表达式的值要在这个范围之内。广义logistic和probit模型形式为:

$$\text{prob}[y = 0] = 1 / (1 + \exp(b_0 + b_1x_1 + b_2x_2 + \dots)) \text{ 及}$$

$$\text{prob}[y = 0] = 1 - p(b_0 + b_1x_1 + b_2x_2 + \dots)$$

这里概率 $p(x)$ 是具有正态分布, 零均值, 方差为1 的随机变量的分布函数。

这两种表达式都使得概率值依赖于观察变量 $x_1, x_2, \dots$  这些变量分别乘以系数 $b_1, b_2$ , 估计模型的目标是找出最好的系数值。当某个特定变量的系数为正值时, 意味着变量的值越高, 对应着 $y=0$  的概率越小, 换句话说, 变量的值越高, 对应于 $y=1$  的概率越大。

广义logistic和probit通常是用最大似然法估计。达到最大值的过程需要迭代方法, 但由于这两种函数的性质很好, 所以在大多数情况下使用平滑方法。

使用LOGIT 命令进行广义logistic估计, LOGIT 要求给出反应变量名和回归变量, 回归变量中要包含常数项。注意命令形式与LS 命令完全相同。但是被解释变量的值只能是0或1。若被解释变量取了其它的观察值, 在估计时对应的样本点要被舍掉。命令输入例子:

```
LOGIT OWNVCR C SIZE INCOME URBAN CABLE
```

对于probit模型, 使用PROBIT命令, 其形式与LOGIT相同, 例如:

```
PROBIT DEMOWIN C GROWTH WAR INFL
```

LOGIT 和PROBIT 命令每一次迭代时都提供了对数似然函数的值, 估计结果包括回归系数、标准差、t-检验值和显著性水平。系数的解释必须要按LOGIT 和PROBIT 函数形式来进行。

此外, 在结果中还包含被解释变量中0 和1 观察值的统计数。

### (三) 系统估计

系统估计是指估计一个完整的联立方程组。在TSP 中使用SYS 命令来完成。SYS 命令联立所有方程有两个主要的优点: 首先, 不同方程的残差经常是互相关联的, SYS 命令允许使用这种互相关性来改善估计的有效性; 其次, 用户可能要限定一个方程的系数与系统中其它一个或多个方程的系数是相同或相关的。这仅仅在联立估计所有方程时才是可能的。

系统估计也有潜在的缺点, 如果错误地说明了一个系统中的一个方程, 在单方程估计中, 仅仅该方程不能正确地估计, 而在系统估计中, 所有的方程都不能被正确地估计。另外, 系统估计方法很费时间, 而且在TSP 中仅能估计70 或少于70 个系数的方程组。但在单方程估计中, 每个方程都可以用到70 个系数。TSP 中, SYS 命令能够估计包含线性或非线性方程的方程组。使用的技术有:

- (0) 普通最小二乘法
- (W) 加权最小二乘法—在方程内取常数权值
- (H) 迭代加权最小二乘法
- (S) 似乎不相关回归
- (I) 迭代似乎不相关回归
- (2) 两阶段最小二乘法
- (J) 加权两阶段最小二乘法
- (3) 三阶段最小二乘法
- (T) 迭代三阶段最小二乘法

在运行SYS 命令后, 若指定了0 或2 选择(普通最小二乘法或两阶段最小二乘法), 则结果在线性方程情况下与LS 或TSLs 估计的结果是完全相同的。唯一的例外出现在对一些方程的系数进行了限制的情况下。非线性方程由修正的NLS 方法估计。这种修正的NLS 比NLS 更加复杂而且需要更长的时间才能达到收敛。

指定了W、H、J 或K 选择进行的加权估计除了对每个方程的标准差的倒数估计值加权外, 与0或2选择是相同的。如果不同的方程有不同的方差, 但不同方程的残差是不相关的,

则这种修正用于处理交叉方程的异方差性。如果没有交叉方程系数限制, 这些技术将给出与0 与2 选择一样的估计结果。然而当存在交叉方程限制时, 它们给出更好的结果。

似乎不相关回归(SUR) 和三阶段最小二乘法(3SLS) 是普通最小二乘法(OLS) 和两阶段最小二乘法(TSLS) 的对应系统方法。由SUR 或3SLS 进行的估计过程是: 首先用单方程估计技术估计方程, OLS 用于SUR, TSLS 用于3SLS。利用第一次迭代的方程系数计算出残差和残差协方差矩阵。在第二次迭代时使用方程残差之间的协方差对第一次估计的系数进行修改。

如果选择了S或3并且系统中的方程都是线性的, 则估计过程在两次迭代之后就完成了。如果选择了I 或T, 或者系统中包含非线性方程, 则在第二次迭代之后重新计算残差和协方差矩阵。估计过程重复下去直到收敛为止。这种技术就是完全信息极大似然法(FIML)。

SYS 的输入是一个系统文件(文本格式), 这个文件可以用TSP 的EDIT 命令或其它的文本编辑软件产生。系统文件中包含要估计的全部方程。除了方程, 系统文件中还可能包含工具变量表和参数初值(仅非线性模型中是需要的)。也可能包含一些定义方程。定义方程并不用在估计过程中但将转换到输出文件中以便于用进一步的模型模拟。

例如, 可以建立一个模型文件MODEL1.SYS 使之包含下述内容:

```
INST LG RD2 AGYG XG XGXY TT4 DCPYP LDCYP LRSL DEP II (-1) S(-1)
```

```
PARAM 1 .5 2 .5 3 1.5
```

```
GG=C(1)+C(2)+II ^ C(3)+(1-C(2))*LG ^ C(3)
```

```
II=(1-C(2))/C(3)+C(4)*GG+C(5)*TT4+C(6)*TT4(-1)+C(7)*LOG(DCPYP)+C(8)*II (-1)
```

```
S=C(20)+C(10)*(GG+TT4)+C(11)*RDZ+C(12)*DEV+C(13)*S(-1)+[AR(1)=C(14)]
```

在文件的第一行中出现的INST 语句指出LG, RD2, ..., DEP 是工具变量。如果没有对每个方程分别给出工具变量表, 则对于2SLS, 加权2SLS, 3SLS 或迭代3LST, INST 语句是必须的。如果选择了OLS、加权LS、SUR 或迭代SUR, INST 语句是不需要的, 此时若说明了INST 语句, 则估计时将其忽略。在每个系统文件中只能包含一个INST 语句。

INST 语句为系统文件中的所有方程说明相同的工具变量表。但是, TSP 不要求所有的方程都使用相同的工具变量表。为了说明某个方程特定的工具变量表, 只需在方程所给出一个“@”符号, 然后指出工具变量表。例如:

```
CONS=C(1)+C(2)*GNP @ C GNP(-1) CONS(-1) M1
```

进行估计时, 方程后面给出的特定工具变量优先于INST语句中给出的工具变量。

PARAM 语句为第一个方程说明参数初值: C(1)=0.5, C(2)=0.5, C(3)=1.5。由于其它的方程关于参数是线性的, 所以没有必要说明参数初值。如果一个方程关于参数是线性的, 则PARAM 语句是无效的。如果对于一个非线性方程没有说明PARAM 语句, 则TSP 就使用系数向量的当前值作为参数初值, 而不管这些当前值是什么。当在系统文件中包含另一个PARAM 语句时, 则后面出现的语句优先用于SYS命令中。

SYS 使用与NLS 同样的方法表述方程。系数依次叫做C(1)、C(2) ... 没有必要连续地使用系数序号。唯一的限制是系数序号不能大于80, 例如C(81) 是非法的, 并且总的系数个数不能超过70 个。

方程既可以是系数非线性也可以是变量非线性或二者均非线性, 方程系数约束可以通过系数运算来实现。例如, 有三个系数出现在三个方程中, 它们的和为1, 则可以这样设定: C(1)在第一个方程中, C(2)在第二个方程中, 然后说明(1-C(1) -C(2)) 作为第三个方程的系数。

一阶自回归误差校正项可以通过在方程后面的方括弧中说明AR(1)=C(常数) 来完成。注意在方括弧中不要含有空格。例如:

$$\text{CONS}=\text{C}(1)+\text{C}(2)*\text{GNP}+[\text{AR}(1)=\text{C}(3)]$$

通过在所有的方程中给出相同的自回归系数序号可以限定系统中的全部方程,也可以给每一个方程说明特定的自回归系数序号。SYS 命令不能估计移动平均误差项和高于一阶的自回归误差校正项。

系统文件中可以包含注释行。注释行以“@”字符开头。例如:

@ 这是注释行

与所有的TSP 命令一样,在用SYS 命令进行估计之前,必须用SMPL 命令说明所使用的样本区间。但是SYS并不需要对所有方程说明相同数目的样本点。如果在SMPL 说明的样本区间中的一部分有一个或多个方程中的序列出现缺值,但在SYS 命令中使用了MD 选择,则SYS 会正确地估计方程系统。

一旦用EDIT 建立了系统文件并用SMPL确定了估计用的样本区间,就为使用SYS 命令做好了准备。完整的SYS 命令格式为:

SYS(m,MD)系统文件名模型文件名

其中各个参数的意义如下:

m 是所使用的估计方法,其值可以为0、W、H、S、2、J、K、3 或T。

MD 选择告诉TSP,即使出现了缺值,估计过程仍继续下去。如果要求TSP 发现缺值样本点时即给出提示信息,请用一个逗号代替MD。

系统文件名:在上面描述过。是包含方程、工具变量表和参数初值的文件的名字。

模型文件名:当用SYS 命令估计一个系统文件时,若想把估计出来的模型输出到一个文件中以便用SOLVE 命令去求解,则要在SYS 命令中给出模型文件名。模型文件名是可选项,若省略了这个名字,则不会生成模型文件。

例如,为了用迭代三阶段最小二乘法估计MODEL1.SYS系统文件并生成MODEL1. DML模型文件,只需输入:

SYS(T) MODEL1.SYS MODEL1.MDL

另一种方法是由TSP提示如何输入。若在统计操作菜单中选择SYS 或直接输入SYS,则TSP提示:

请输入系统文件名?

回答MODEL1.SYS。若想把估计结果输出到一个文件中,则可回答:

MODEL1.SYS MODEL1.MDL

这样MODEL1.MDL 中将包含估计出的模型。然后TSP 提问:

- (0) 普通最小二乘法
- (W) 加权最小二乘法—在方程内取常数权值
- (H) 迭代加权最小二乘法
- (S) 似乎不相关回归
- (I) 迭代似乎不相关回归
- (2) 两阶段最小二乘法
- (J) 加权两阶段最小二乘法
- (3) 三阶段最小二乘法
- (T) 迭代三阶段最小二乘法

由于MODEL1 中包含关于参数非线性的方程,所以选择S 和I 是一样的,选择3 和T 也是一样的。为了由三阶段最小二乘法进行估计,请选择T。

在系统文件估计完成后,可以通过输入:

表10.1 1749—1924年太阳黑子年平均数(176个数据)

80.9	83.4	47.7	47.8	30.7	12.2	9.6	10.2	32.4	47.6
54.0	62.9	85.9	61.2	45.1	36.4	20.9	11.4	37.8	69.8
106.1	100.8	81.6	66.5	34.8	30.8	7.0	19.8	92.5	154.4
125.9	84.8	68.1	38.5	22.8	10.2	24.1	82.9	132.0	130.9
118.1	89.9	66.6	60.0	46.9	41.0	21.3	16.0	6.4	4.1
6.8	14.5	34.0	45.0	43.1	47.5	42.2	28.1	10.1	8.1
2.5	0.0	1.4	5.0	12.2	13.9	35.4	45.8	41.1	30.4
23.9	15.7	6.6	4.0	1.8	8.5	16.6	36.3	49.7	62.5
67.0	71.0	47.8	27.5	8.5	13.2	56.9	121.5	138.3	103.2
85.8	63.2	36.8	24.2	10.7	15.0	40.1	61.5	98.5	124.3
95.9	66.5	64.5	54.2	39.0	20.8	6.7	4.3	22.8	54.8
93.8	95.7	77.2	59.1	44.0	47.0	30.5	16.3	7.3	37.3
73.9	139.1	111.2	101.7	66.3	44.7	17.1	11.3	12.3	3.4
6.0	32.3	54.3	59.7	63.7	63.5	52.2	25.4	13.1	6.8
6.3	7.1	35.6	73.0	84.9	78.0	64.0	41.8	26.2	26.7
12.1	9.5	2.7	5.0	24.4	42.0	63.5	53.8	62.0	48.5
43.9	18.6	5.7	3.6	1.4	9.6	47.4	57.1	103.9	80.6
63.6	37.6	26.1	14.2	5.8	16.7				

SOLVE MODEL1.MDL

来运行动态模型(提供的模型已存储在MODEL1.MDL 中)。

## §12.3 用例与样本程序

### §12.3.1 用例

【例12.1】时间序列模型分析中,太阳黑子数据是比较经典的例子,现对它进行分析:设上述数据存于文件S.TXT,读取数据的MicroTSP命令为:

```
>creat a 1749 1924
>read(s) s.txt sunspot
```

根据Pandit和吴贤铭的建模策略,首先拟合ARMA(2,1)和ARMA(4,3)模型:

```
>ls sunspot c ma(1) ar(1) ar(2)
```

结果为:

```
LS // Dependent Variable is SUNSPOT
Date: 12-14-1993 / Time: 14:33
SMPL range: 1751 - 1924
Number of observations: 174
Convergence achieved after 4 iterations
```

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	44.480818	3.9717037	11.199430	0.000
MA(1)	-0.1614766	0.1171300	-1.3786103	0.168
AR(1)	1.4284049	0.0884517	16.148989	0.000
AR(2)	-0.7245218	0.0793022	-9.1362137	0.000
R-squared	0.804150	Mean of dependent var	44.33736	
Adjusted R-squared	0.800694	S.D. of dependent var	34.73446	
S.E. of regression	15.50675	Sum of squared resid	40878.08	
Durbin-Watson stat	1.979609	F-statistic	232.6706	
Log likelihood	-721.8539			

```

>ls sunspot c ma(1) ma(2) ma(3) ar(1) ar(2) ar(3) ar(4)

```

结果为:

```

LS // Dependent Variable is SUNSPOT
Date: 12-14-1993 / Time: 14:35
SMPL range: 1753 - 1924
Number of observations: 172
Convergence achieved after 10 iterations

```

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	44.671449	1.5564850	28.700212	0.000
MA(1)	0.9301516	0.8837230	1.0525375	0.293
MA(2)	0.3950168	0.5927001	0.6664700	0.505
MA(3)	-0.1018366	0.1219219	-0.8352611	0.404
AR(1)	0.3676083	0.8801520	0.4176646	0.676
AR(2)	0.2755596	0.7192365	0.3831280	0.702
AR(3)	-0.0348496	0.4872030	-0.0715299	0.943
AR(4)	-0.3663540	0.4745024	-0.7720804	0.440
R-squared	0.812607	Mean of dependent var	44.29767	
Adjusted R-squared	0.804608	S.D. of dependent var	34.93502	
S.E. of regression	15.44237	Sum of squared resid	39108.53	
Durbin-Watson stat	1.999424	F-statistic	101.5951	
Log likelihood	-710.7452			

由于ARMA(2,1)和ARMA(4,3)的残差平方和(Sum of squared resid) 分别为SSRI=40878.08和SSRII=39108.53, 故F 统计量为:

$$F = [(40878.08 - 39108.53) / 4] / [39108.53 / (176 - 8)] < F_{0.95}(4, \infty) = 2.37$$

表明无需拟合比ARMA(2,1)更高阶的模型。注意到在ARMA(2,1)中, MA(1) 的参数偏小。且其置信区间包含0, 故F下面考虑AR(2)模型。

```

>ls sunspot c ar(1) ar(2)

```

结果为:

```

LS // Dependent Variable is SUNSPOT

```

```

Date: 12-14-1993 / Time: 14:37
SMPL range: 1751 - 1924
Number of observations: 174
Convergence achieved after 1 iterations
      VARIABLE      COEFFICIENT      STD. ERROR      T-STAT.      2-TAIL SIG.
      C              44.393596       3.7546702       11.823567       0.000
      AR(1)          1.3361112       0.0580720       23.007858       0.000
      AR(2)          -0.6501527      0.0581050       -11.189270      0.000
R-squared              0.801962      Mean of dependent var  44.33736
Adjusted R-squared    0.799646      S.D. of dependent var  34.73446
S.E. of regression    15.54746      Sum of squared resid   41334.70
Durbin-Watson stat    2.121557      F-statistic           346.2362
Log likelihood         -722.8203

```

比较AR(2)与ARMA(2,1), 其F统计量值为:

$$F = [(41334.78 - 40878.08) / 1] / [40878.08 / (176 - 4)] < F_{0.95}(1, \infty) = 3.84$$

因此, 从上面F值知AR(2)对太阳黑子活动数据是合适的。

【例12.2】非线性回归用例(王学仁等编《应用回归分析》, 重庆人民出版社):

设线性回归方程为:  $Y = \alpha + (0.49 - \alpha) \exp[-\beta(x - 8)]$ , 观测数据(X,Y)为:

10 .48

20 .42

30 .40

40 .39

设数据存于文件NL.DAT, TSP 分析程序NL.PRM为:

```

create (u) 4
read(o) nl.dat x y
nls y=c(1)+(0.49-c(1))*exp(-c(2)*(x-8))
param 1 0 2 1
option output c:nl.out
pon

```

在TSP的";"提示符下打入RUN NL.PRM 命令, 经九次迭代, 结果:

```

C(1)  0.403279  C(2)  0.509604  SSR  0.00252374
C(1)  0.406802  C(2)  0.161704  SSR  0.00055278
C(1)  0.397746  C(2)  0.102024  SSR  0.00025669
C(1)  0.389350  C(2)  0.096182  SSR  7.7215E-05
C(1)  0.384984  C(2)  0.086886  SSR  5.2099E-05
C(1)  0.382316  C(2)  0.082154  SSR  4.6129E-05
C(1)  0.381111  C(2)  0.080126  SSR  4.5305E-05
C(1)  0.380774  C(2)  0.079565  SSR  4.5257E-05
C(1)  0.380730  C(2)  0.079492  SSR  4.5257E-05

```

Convergence achieved after 9 iterations

```
=====
                COEFFICIENT   STD. ERROR   T-STAT.   2-TAIL SIG.
=====
                C(1)         0.3807300   0.0085770   44.389472   0.001
                C(2)         0.0794917   0.0162171   4.9017227   0.039
=====
R-squared                0.990717   Mean of dependent var   0.422500
Adjusted R-squared       0.986075   S.D. of dependent var   0.040311
S.E. of regression       0.004757   Sum of squared resid    4.53E-05
Durbin-Watson stat      1.958976   F-statistic              213.4379
Log likelihood           17.10316
=====
```

即  $\alpha = C(1) = 0.3807, \beta = C(2) = 0.07949$ 。

【例12.3】多方程系统估计。仍用第4章食品消费和价格的例子，系统方程为：

$$Y = C(10) + C(11) X_1 + C(12) X_2$$

$$Y = C(20) + C(21) X_1 + C(23) X_3 + C(24) X_4$$

原始数据(Y,X1,X2,X3,X4)为：

```
98.485  100.323  87.4  98  1
.....
106.232 113.49  127.1  93  20
```

设数据存于文件ECO.DAT中，TSP 程序ECO.PRM为：

```
create a 1 20
read(o) eco.dat y x1 x2 x3 x4
option output e:eco.out
pon
sys(o) ecomd.sys
```

模型ECOMD.SYS为：

```
inst x1 x2 x3 x4
y=c(10)+c(11)*x1+c(12)*x2
y=c(20)+c(21)*x1+c(23)*x3+c(24)*x4
```

在TSP” > ”提示符下运行命令RUN ECO.PRM，完成参数估计。注意：

1. 对普通最小二乘不必指定工具变量。
2. 对两步最小二乘和三步最小二乘，程序中指定工具变量为：X1,X2,X3,X4。
3. 在上述程序中在采用普通最小二乘、两步最小二乘和三步最小二乘估计时，语句：SYS(\*)

ECOMD.SYS 中的星号(\*)应分别换为0,2,3。

输出结果如下表：

```
=====
                OLS          COEFFICIENT   STD. ERROR   T-STAT.   2-TAIL SIG.
=====
```

=====				
C(10)	99.839109	7.4941890	13.322203	0.000
C(11)	-0.3160949	0.0903416	-3.4988839	0.003
C(12)	0.3350539	0.0452963	7.3969450	0.000
=====				
C(20)	58.276931	11.461841	5.0844304	0.000
C(21)	0.1603437	0.0948678	1.6901806	0.110
C(23)	0.2481402	0.0461884	5.3723480	0.000
C(24)	0.2483209	0.0975180	2.5464119	0.022
=====				
=====				
2LS	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
=====				
C(10)	99.839083	7.4941789	13.322218	0.000
C(11)	-0.3160940	0.0903415	-3.4988790	0.003
C(12)	0.3350533	0.0452963	7.3969295	0.000
=====				
C(20)	58.276726	11.461841	5.0844123	0.000
C(21)	0.1603444	0.0948677	1.6901901	0.110
C(23)	0.2481415	0.0461884	5.3723729	0.000
C(24)	0.2483215	0.0975180	2.5464184	0.022
=====				
=====				
3LS	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
=====				
C(10)	99.215728	6.9048101	14.369074	0.000
C(11)	-0.2703525	0.0812652	-3.3267929	0.004
C(12)	0.2945319	0.0385021	7.6497632	0.000
=====				
C(20)	62.364841	9.9066193	6.2952697	0.000
C(21)	0.1459674	0.0844489	1.7284697	0.103
C(23)	0.2117410	0.0356311	5.9425863	0.000
C(24)	0.3308977	0.0606537	5.4555216	0.000
=====				

模型输出R-平方、调整R-平方、回归标准误、Durbin-Watson 统计量、因变量的均值与标准差，均方误差和、F-统计量，此处从略。

### §12.3.2 样本程序

MicroTSP 6.5 系统同时附加了样本演示程序，对该软件的掌握很有帮助，现介绍如下。

★第一个样本程序，是想预测Holiday Inns hotel 的收入，演示数据处理、回归和预测。第一步是建立一个工作文件，它保存于计算机的RAM中，这可以由CREATE命令完成。使用该命令后，系统提问数据的频度，时间的区间，现在是1970 到1981的年度数据。也可以直接使用命令：

```
CREATE A 70 81
```

下一步使用数据编辑器来录入数据。序列名至多八个字符，使用DATA命令或在数据操作菜单下选择进入数据编辑。MicroTSP提示要编辑的序列名，然后录入数据：

```
697
814
963
1122
1224
1369
1569
```

假设应该录入的数据是1539，我们可以用命令.B返回去把它修正，继续录入：

```
1780
2161
2605
2915
```

最后一个观察是1980 的，因此，我们完成了数据录入，使用命令X退出。现在可以看出REV出现在状态窗口了。

也许我们应该把序列REV 存到磁盘上，当需要时FETChed 调入。

```
> STORE REV
```

要在数据库文件中包括说明，可以使用LABEL 命令插入或调出。

```
> LABEL REV Total room revenue, in millions of dollars
```

当LABEL不带其它参数时，现有的信息就会出现在屏幕上，如：

```
> LABEL REV
```

为进行统计分析，还需要一些其它序列的数据，现使用FETCh命令。先用DIR命令浏览一下文件：

```
> DIR
```

所有数据文件都用扩展名.DB, 因此可直接使用DIR \*.DB 命令来浏览。结合上面的LABEL命令可以看出每个文件的描述。如对GNP.DB, 使用命令：

```
> LABEL GNP
```

```
> FETCh GNP OCCUP
```

现在列出磁盘上可能对分析Holiday Inns 数据有用的序列的名字。

REV Total room revenue, in millions of dollars  
 OCCUP Occupancy rate - percent of rooms occupied on the average night  
 RRATE Room rate - average revenue per occupied room per night  
 ROOMS Number of rooms in the Holiday Inns system  
 GNP U.S. Gross National Product in current dollars  
 PGNP Implicit price deflator for GNP  
 UNEMP U.S. Unemployment rate  
 CPR Interest rate on prime commercial paper

得到其它的序列:

```
> FETCH RRATE ROOMS GNP PGNP UNEMP CPR
```

现在可以浏览我们的数据了, 首先应告知数据的范围。使用SMPL 命令指定:

```
> SMPL 70 80 使用SHOW 命令得到数据的列表。
```

```
> SHOW REV OCCUP RRATE ROOMS GNP UNEMP
```

MicroTSP 也允许用户以图形方式浏览数据, 在第二个演示程序中说明。

通常, 数据在进行统计分析之前要进行转换, GENR 就是这样的命令。如产生序列RATIO, 可以使用命令:

```
> GENR RATIO=REV/GNP
```

MicroTSP 处理数据滞后也很方便。首先, 改变观察的范围:

```
> SMPL 71 80
```

产生一个序列, 使它等于RATIO 变化的百分比, 使用命令:

```
> GENR PCHR=100*(RATIO-RATIO(-1))/RATIO(-1)
```

GENR 命令也可用于动态模拟, 以及对数据序列施加过滤。一个简单的用例是产生一个趋势变量, 取值为0 in 1970, 1 in 1971, 2 in 1972,....., 11 in 1981。

```
> SMPL 70 70
```

```
> GENR TREND=0
```

```
> SMPL 71 81
```

```
> GENR TREND=TREND(-1)+1
```

事实上, MicroTSP 中有更为方便的方法, 产生一个从1970 取值为零开始的趋势变量, 使用命令:

```
> GENR TREND=@TREND(70)
```

```
> SHOW RATIO PCHR TREND
```

现在可以对上述数据进行一些统计分析了, MicroTSP 把最重要的结果先列出来, 然后询问用户是否需要得更为详细, 是在打印机(P)还是在屏幕(S), 或者跳过去(Enter)。

首先可以浏览数据的一些描述统计量: 均值、标准差、最大值、最小值、协方差和相关。

```
> SMPL 71 80
```

```
> COVA REV OCCUP GNP UNEMP
```

在屏幕的最下一行, MicroTSP 会询问我们是否重复COVA 的输出。

数据的检查之后, 我们可以产生关于REV 的预测方程, 构造REV 关于常数(C)、上年的国民收入GNP(-1)和收入REV(-1)。C 由系统设定, 取值皆 1。简单的最小二乘回归命令是LS, 格式是:

```
> LS REV C GNP(-1) REV(-1)
```

MicroTSP 继续给出一些提示, 与“S”和“P”相应, 提示“G”表示高分辨率图形显示。MicroTSP 也把最小二乘的结果保留在内存中, 这些结果可用于预测。使用命令SHOWEQ 命令可以浏览内存的模型。

```
> SHOWEQ
```

现要指明预测的范围, 如仅仅是1981年的情况:

```
> SMPL 81 81
```

则可以使用命令FORCST, 它带一个参数用于存贮预测的序列REVF, REVF在19 81年以前的值就是REV的值。

```
> FORCST REVF
```

最后可以统览一下:

```
> SMPL 70 81
```

```
> SHOW REV REVF
```

```
> SMPL 71 80
```

存贮工作文件, 使用系统提供的SAVE命令, 以后用LOAD就可以调入了。

```
> SAVE HOLIDAY
```

★MicroTSP 有许多命令用于产生高分辨图形, 现在就用上面生成的工作文件HOLIDAY。

```
> LOAD HOLIDAY
```

由于REV和GNP是相关联的, 这个关系可以通过图形来体现。MicroTSP 支持几种图形格式: 线图、散点图、条图、圆图以及条图和线图的混合。图形可以在数据管理菜单(F4)下选择, PLOT 命令对应的是线图, MicroTSP 为线图提供了尺度选择, 但我们可以用的它自动选择。

```
> PLOT(A) REV GNP
```

屏幕下方的一些选择, 选择T (Type) 允许在屏幕上给出标号, 用P (Print)则打印图形, S (Save) 则贮一个MicroTSP 图形文件, O(Options) 则允许用户改变图形, 如图形的宽度、边界、字型、尺度等。

同样的数据可以用条图和散点图表达。

```
> BAR REV GNP
```

```
> SCAT REV GNP
```

★第三部分进一步演示MicroTSP 时序分析。假设用户对于Box-Jenkins 分析、两阶段最小二乘法、多项分布滞后是熟悉的。用MicroTSP构造一个汽车销售情况的简单模型, 变量是:

```
TBQ      3-month Treasury bill rate
```

```
ASQ      automobile sales in millions at annual rates
```

使用CREATE 产生一个季节型的工作文件。用特别的记号, 如1958. 1 表示1958 年第一季度的数据。

```
> CREATE Q 58.1 78.4
```

```
> FETCH TBQ ASQ
```

```
> GENR TIME=@TREND(58.1)
```

首先产生一个ASQ 的单变量的Box-Jenkins 模型, 把ASQ作为一个ARMA过程。为便于正确建模, 首先看一下ASQ的相关图和偏相关图, 由IDENT命令来完成。如浏览16个季节的自相关的偏自相关, 打入命令:

```
> IDENT(16) ASQ
```

从相关图可以猜想, ASQ 是一个ARMA(2,0) 过程。LS 命令将估计这个ARMA 模型, 由于ASQ具有正的均值, 指定C作为常数项, 但这里不是外生变量, 而是AR() 和MA()的项, 对1960.3 到1978.4 数据拟合模型, 我们打入:

```
> SMPL 60.3 78.4
```

```
> LS ASQ C AR(1) AR(2)
```

现在可以看一下模型的新息(残差)的相关图, 是否还存在自相关, MicroTSP 使用变量RESID 存贮新近估计的残差。

```
> IDENT(16) RESID
```

在滞后4 和8 处仍有一些相关, 提示季节项的存在, 现使用MA(4) 和MA(8) 项:

```
> LS ASQ C AR(1) AR(2) MA(4) MA(8)
```

虽然可以继续使用单变量的ARMA 模型, 但我们想到利率会对销售额有一定的影响。可以对ASQ关于TBQ估计一个简单线性回归模型, 然而模型将有序列相关, 但是ARMAX技术使我们可以使用广义最小二乘法估计。

```
> LS ASQ C TIME TBQ TBQ(-1) TBQ(-4) AR(1) AR(2) MA(4)
```

看来TBQ 的确有益于预测汽车的销售情况。注意到TBQ有一个没有预料的正的系数, 而滞后的TBQ则是负号。还可以使用SHOWEQ 命令显示模型的估计情况。

模型的系数也可用于GENR的计算, 如:

```
GENR ASHAT = C(1)+C(2)*TBQ+C(3)*TBQ(-1)+C(4)*TBQ(-4)
```

给定利率解释销售的重要性, 我们可以在ASQ方程中包括更多的TBQ滞后项, 但可能会有共线性的问题。处理这一问题的一种方法是限制TBQ系数在多项式上, 这用到了Polynomial Distributed Lags技术。使用这一技术时, MicroTSP 建立临时性序列PDL1,PDL2,... 用于估计。现对于ASQ, 使TBQ 的八个滞后限制于三阶多项式, 使用命令:

```
> LS ASQ C TIME PDL(TBQ,8,3) AR(1) AR(2) MA(4)
```

看来结果并没有改善多少, 所以仍然回到原来的模型。TBQ 的系数仍然是正的, 这可能由于ASQ与TBQ关联的结果。高的汽车需求可能对利率施加压力, 现暂时忽略ARMA误差, 试一下ASQ的两阶段最小二乘方程, 这可以由位置符号@引导回归因子和工具变量来完成。

```
> TSLS ASQ C TIME TBQ TBQ(-1) TBQ(-4) @ C TIME TBQ(-1) TBQ(-2) TBQ(-3) TBQ(-4) ASQ(-1)
```

本例修正了simultaneity对TBQ系数的改变影响很小。

★第四个样本程序介绍一些更高级的技术, 即非线性模型、probit 和二分类logit 模型。在计量经济学中有两种非线性, 即变量或其系数的非线性, 变量非线性而参数为线性则可由LS 或TSLS估计, 如模型:

$$Y = a + b * \text{LOG}(L) + c * \text{LOG}(K)$$

关于L 和K 是非线性的, 但其参数是线性的, 可以首先创建变量, 然后用普通的最小二乘。

```
> GENR LOGL = LOG(L)
```

```
> GENR LOGK = LOG(K)
```

```
> LS Y C LOGL LOGK
```

这种做法不适于参数是非线性的情况, 如CES production function 模型:

$$Q = a + b * \text{LOG}(c * L^d + (1 - c) * K^d)$$

Q,L, 和K, 是对数log output, labor and capital。a,b,c 和d 是系数。

这时应使用非线性最小二乘命令NLS, 它使用特殊的写法, 其参数记为C(1), C(2), 等等。针对上述模型有:

```
> NLSQ = C(1) + C(2) * LOG(C(3) * L^C(4) + (1 - C(3)) * K^C(4))
```

使用NLS 之前, 应提供参数的初值, 这可以由PARAM完成, 本例就是:

```
> PARAM 1 .5 2 -1 3 .5 4 -1
```

结果C(1) 为.5, C(2) 为-1, 等等。

现看此命令的用法, 首先产生一个undated工作文件。

```
> CREATE U 30
```

把序列取到工作文件:

```
> FETCH Q L K
```

其次估计CES 函数。

```
> PARAM 1 .5 2 -1 3 .5 4 -1
```

```
> NLS Q=C(1)+ C(2)*LOG(C(3)*L^C(4)+(1-C(3))*K^C(4))
```

同线性估计一样, MicroTSP 也把方程存于内存, 可以使用SHOWEQ 命令浏览。使用STOREQ 命令则可以把方程存于磁盘。使用FORCST 或FIT 可以进行预测。

现在转向probit 和logit 模型, 一些人在工作训练中成功与否, 受年龄、智力、文化的影响, logit 和probit 模型形式是:

$$\text{Prob}[\text{PASS} = 0] = 1 / (1 + \exp(a + b * \text{AGE} + c * \text{IQ} + d * \text{EDUC}))$$

$$\text{Prob}[\text{PASS} = 0] = 1 - P(a + b * \text{AGE} + c * \text{IQ} + d * \text{EDUC})$$

仍然产生一工作文件, 取序列并估计:

```
> CREATE U 38
```

```
> YFETCH PASS AGE EDUC IQ
```

```
> LOGIT PASS C AGE EDUC IQ
```

```
> PROBIT PASS C AGE EDUC IQ
```

```
> SHOWEQ
```

继续可使用FORCST, FIT, 或SOLVE。

★最后一个例子是考察portfolio of Wells Fargo Bank 并给出两阶段最小二乘和系统估计示例。数据是1963—1983年间Wells Fargo 的年度数据以及全国和加州地区在这个期间的经济情况。序列是:

SECU Wells Fargo's holdings of securities, in millions of dollars

LOAN Their holdings of loans

RSECU Percentage return earned by Wells Fargo on securities

RLOAN Percentage return on loans

SALES Retail sales in California, in millions of dollars

RCP6M Interest rate on 6-month commercial paper, percent

CONTR Value of construction contracts in California, millions of dollars

TREND Time trend, starting with a 1 in 1963

生成工作文件: PORT 是银行全部portfolio的大小, SPREAD是percentage spread between the interest earned on loans and the 6-month commercial paper rate。

由于RLOAN序列始于1964, MicroTSP不能计算SPREAD在1963年的值, 因此赋值NA(not available)。

LOARAT是portfolio held in loans的比例。

LOASAL是Wells Fargo's 贷款与销售的比。

CONSAL是建设与零售的比。

```
> CREATE A 63 83
> FETCH SECU LOAN RSECU RLOAN SALES RCP6M CONTR TREND
> GENR PORT=LOAN+SECU
> GENR SPREAD=RLOAN-RCP6M
> GENR LOARAT=LOAN/PORT
> GENR LOASAL=LOAN/SALES
> GENR CONSAL=CONTR/SALES
```

分析的目的是了解银行从贷款获利的比例以及用贷款保持的portfolio比例。应该考虑: 贷款的提供与贷款利率成正相关, 为使顾客购买更旺, 银行应使贷款利率相对于其它方面更加诱人, 贷款需求与贷款利率成负相关。因此我们分析的目的是估计贷款的供求方程, 以分出这两种影响。

对供给一方市场, 我们对LOARAT建模( fraction of Wells Fargo's portfolio held in loans), 假设Wells Fargo在时返还贷款增加时增加贷款而在其它securities相对高时减少贷款。因而用LOARAT对返还贷款RLOAN、其它securities RSECU及LOARAT滞后行回归。RSECU, LOARAT(-1), LOASAL(-1) CONSAL, PORT, SALES, 和TREND作为工具变量。

TSLs命令是:

```
> SMPL 64 83
> TSLs LOARAT C RLOAN RSECU LOARAT(-1) @
      C RSECU LOARAT(-1) LOASAL(-1) CONSAL PORT SALES TREND
```

估计的供给方程应予存贮。

```
> STOREQ SUPPLY
```

现看需求方程。使用贷款率、RLOAN和其它一些主要的率的差很方便, 该差是SPREAD。我们假设高的贷款需求源于高的零售, 建设增加贷款利率和其它securities间的差。因此SPREAD的右端是LOASAL, 贷款对零售的比、LOASAL的滞后、建设与零售的比CONSAL以及TREND。工具变量是LOASAL(-1), CONSAL, and TREND plus RSECU, SALES, and PORT.

TSLs命令是:

```
TSLs SPREAD TREND LOASAL LOASAL(1) CONSAL(-1) RSECU
      LOARAT(-1) LOASAL(-1) CONSAL PORT SALES TREND
> STOREQ DEMAND
```

现在应转到完整的模型上来, 第一个方程告知银行如何根据贷款和securities利率作出决定, 第二个方程告知应该如何让顾客有更多的购买力, 完整就是把两者结合起来。MicroTSP中, 要求编辑一个模型文件, 其包括模型的方程和其它有关信息。它有三种语句, 即ASSIGN(告知MicroTSP求解的序列存于不同的名)、标识量(identity)、估计方程, 本例为:

```
ASSIGN LOARAT PLORAT LOAN PLOAN SPREAD PSPRED
ASSIGN RLOAN PRLOAN LOASAL PLOASA
```

```

LOAN = LOARAT*PORT
RLOAN = SPREAD + RCP6M
LOASAL = LOAN/SALES
LOARAT = .2463962 +2.786724E-02*RLOAN -2.735787E-02*RSECU
        +.6360336*LOARAT(-1)
SPREAD = -4.277588 +.1052442*TREND -112.2502*LOASAL
        +103.5918*LOASAL(-1) +20.45162*CONSAL

```

可以使用MicroTSP的EDIT 编写模型:

```
> EDIT LOAN
```

EDIT 有许多点命令, 使用.F 命令取得估计的方程。

```
.F SUPPLY
```

```
.F DEMAND
```

使用.L 命令浏览模型, 使用.X 退出编辑。

使用TYPE LOAN命令可以证实模型确已存贮。

现用原始外生变量得到一个基础解。

```
> SMPL 76 83
```

```
> SOLVE LOAN
```

```
> GENR PSECU = PORT - PLOAN
```

```
> SHOW LOAN PLOAN SECU PSECU RLOAN PRLOAN
```

下一步是用模型模拟贷款市场, 看建设是否在1976到1983要高10%。

```
> GENR CONSAL = 1.1*CONSAL
```

我们可以用EDIT 编辑一个与LOAN 类似的模型SIMLN, 只要把前缀P改成S 即可。事实上软件已经事先准备好了, 只消浏览一下。

```
> TYPE SIMLN
```

```
> SOLVE SIMLN
```

```
> GENR SSECU = PORT - SLOAN
```

```
> SHOW PLOAN SLOAN PSECU SSECU PRLOAN SRLOAN
```

仍然把CONSAL 恢复过来, 并且看在贷款市场改变供给的影响。

```
> GENR CONSAL = CONSAL/1.1
```

```
> GENR PORT = PORT*1.1
```

```
> SOLVE SIMLN
```

```
> GENR SSECU = PORT - SLOAN
```

```
> SHOW PLOAN SLOAN PSECU SSECU PRLOAN SRLOAN
```

```
> GENR PORT = PORT/1.1
```

上述样本程序没有演示以下功能:

SYS	系统估计, 包括三阶段最小二乘和似乎不相关回归
VAREST	向量自回归估计。
VARSTAT	从VAR 模型获得冲激响应函数及方差分解。
SAVE and LOAD	存贮工作文件。
EXPAND	改变工作文件的大小。
LGRAPH and PGRAPH	调入与打印磁盘图形。
READ and WRITE	与其它程序的数据交换(包括Lotus 123 工作表及 PRN 文件, WordStar 文件, DIF 文件) 及读取大型计算机而来的数据。
HIST	直方图。
CROSS	产生互相关图(cross correlograms)。
SMA and SAR	在ARMA和ARMAX模型引入季节(multiplicative) 移动平均及自回归误差项。
SMOOTH	指数平滑(single, double, Holt-Winters additive & multiplicative)
SEAS	季节调整。
STOREQ and FETEQ	存贮估计的方程。
SORT	按照一个或多个关键字对工作文件排序。
CONV	改变时间序列的频度, 如四季资料到年份资料。
PON,POFF & TRACE	打印机控制。
RUN	批处理方式运行MicroTSP程序。
REN and DEL	磁盘工作文件换名和删除。
R and D	在工作文件中的序列更名和删除。
RND and NRND	产生随机数(GENR & models)。
OPTION	设定MicroTSP选项, 包括输出再定向、屏幕和功能键控制等。

其它的特色, 如:

- 统计输出和数据表格可转成磁盘文件或打印。
- 批处理操作, 允许用户用哑元传递书写MicroTSP程序, 从而允许宏操作。
- 这些程序可以包括FOR..NEXT循环。
- 在SHOW 和PRINT 中使用用户提供的格式输出数据表。
- 在PLOT 命令中定义输出格式。
- EDIT 可以编辑程序的模型文件。
- SOLVE 可以处理具有ARMA 误差的模型。