第3次作业

MATLAB

```
clear;
close all;
clc;

tutu = load('earn.txt');

EXP = tutu(:,1);
EXP2 = tutu(:,2);
WKS = tutu(:,3);
OCC = tutu(:,4);
IND = tutu(:,5);
SOUTH = tutu(:,6);
SMSA = tutu(:,7);
MS = tutu(:,8);
FEM = tutu(:,9);
UNION = tutu(:,10);
ED = tutu(:,11);
BLK  = tutu(:,12);
LWAGE = tutu(:,13);
M = tutu(:,14);
F_EDC= tutu(:,15);

Y = UNION;
N = length(Y);
X = [EXP, WKS, OCC, IND, SOUTH, SMSA, MS, FEM, ED, BLK, ones(N,1)];
K = size(X,2);

%用内置函数fitlm完成回归
tbl = table(EXP, WKS, OCC, IND, SOUTH, SMSA, MS, FEM, ED, BLK, UNION, 'VariableNames', ...
    {'EXP', 'WKS', 'OCC', 'IND', 'SOUTH', 'SMSA', 'MS', 'FEM', 'ED', 'BLK', 'UNION'});
mdl = fitlm(tbl, 'UNION~EXP+WKS+OCC+IND+SOUTH+SMSA+MS+FEM+ED+BLK');

%手写回归
betaHat = X'*X \ X'*Y;  %估计量
uHat = Y - X*betaHat;  %离差
sigma2Hat = sum(uHat.^2) / (N-K);  %随机误差项方差估计值
varBetaHat = sigma2Hat * inv(X'*X);  %方差-协方差矩阵
stdBetaHat = sqrt(diag(varBetaHat));  %标准误
t = (betaHat-zeros(K, 1)) ./ stdBetaHat;
pValue = 2*(1-normcdf(abs(t)));

mdl
betaHat = betaHat'
t = t'
```

fitlm 函数结果：

线性回归模型:

UNION ~ 1 + EXP + WKS + OCC + IND + SOUTH + SMSA + MS + FEM + ED + BLK

估计系数:

	Estimate	SE	tStat	pValue
(Intercept)	1.1959	0.22701	5.2679	1.9433e-07
EXP	-0.0019737	0.0017264	-1.1433	0.25339
WKS	-0.017809	0.0034187	-5.2092	2.6329e-07
OCC	0.31812	0.046425	6.8523	1.8501e-11
IND	0.030048	0.038072	0.78923	0.4303
SOUTH	-0.17013	0.039801	-4.2745	2.2374e-05
SMSA	0.084522	0.038464	2.1974	0.028382
MS	0.098953	0.063781	1.5515	0.12133
FEM	-0.10871	0.079266	-1.3714	0.17078
ED	-0.016187	0.0085923	-1.8839	0.060072
BLK	0.050197	0.07113	0.70571	0.48065

观测值数目: 595，误差自由度: 584
均方根误差: 0.426
R 方: 0.234，调整 R 方 0.22
F 统计量(常量模型): 17.8，p 值 = 1.81e-28

手写结果:

```
betaHat =

  列 1 至 10

   -0.0020   -0.0178    0.3181    0.0300   -0.1701    0.0845    0.0990   -0.1087   -0.0162    0.0502

  列 11

    1.1959


t =

  列 1 至 10

   -1.1433   -5.2092    6.8523    0.7892   -4.2745    2.1974    1.5515   -1.3714   -1.8839    0.7057

  列 11

    5.2679
```