Robust_estimator

zhao hongqiang

5/24/2021

Robust estimator of a parameter of the probability distribution

data choise Percentage changes in quarterly personal consumption expenditure, personal disposable income, production, savings and the unemployment rate for the US, 1960 to 2016.

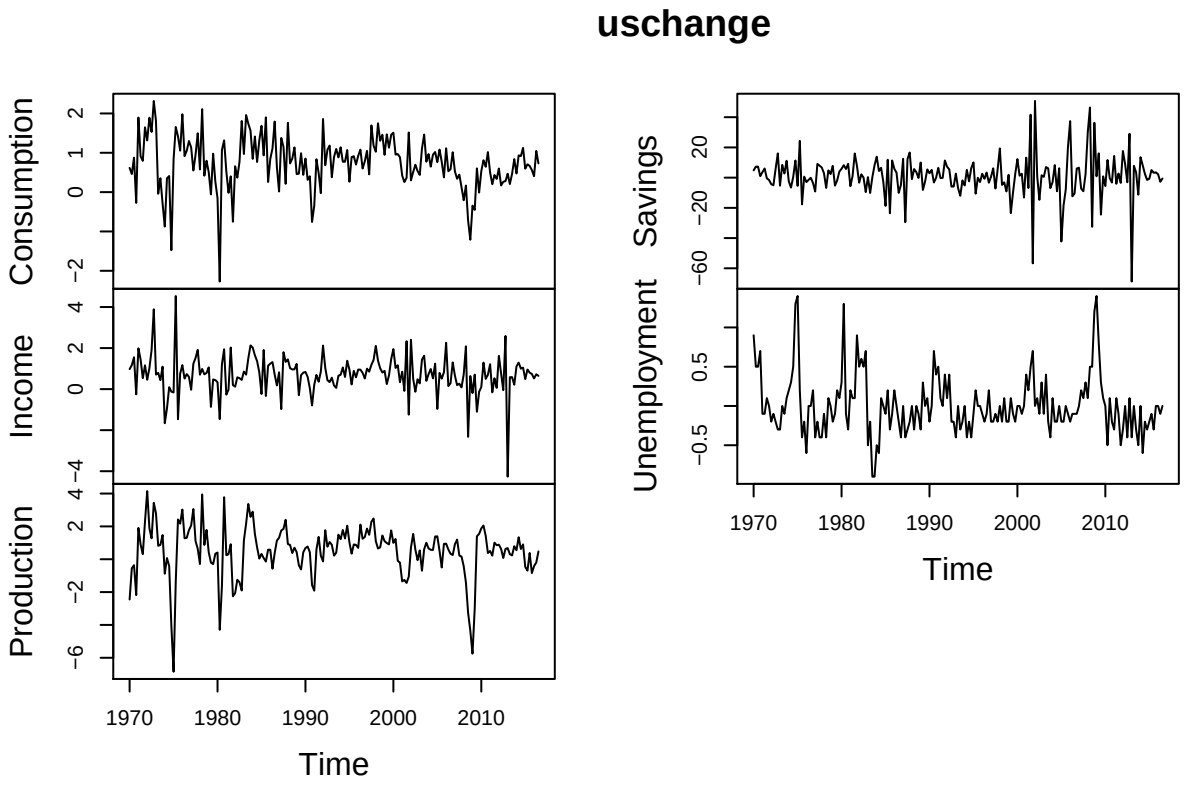
```
summary(uschange)
```

```
##   Consumption          Income          Production          Savings       
##  Min.   :-2.2741   Min.   :-4.2652   Min.   :-6.85104   Min.   :-68.788  
##  1st Qu.: 0.4198   1st Qu.: 0.3378   1st Qu.: 0.05568   1st Qu.: -4.218  
##  Median : 0.7721   Median : 0.7237   Median : 0.65793   Median :  1.280  
##  Mean   : 0.7465   Mean   : 0.7176   Mean   : 0.50806   Mean   :  1.222  
##  3rd Qu.: 1.0898   3rd Qu.: 1.1650   3rd Qu.: 1.30572   3rd Qu.:  6.651  
##  Max.   : 2.3183   Max.   : 4.5365   Max.   : 4.14957   Max.   : 50.758  
##   Unemployment      
##  Min.   :-0.900000  
##  1st Qu.:-0.200000  
##  Median : 0.000000  
##  Mean   : 0.007487  
##  3rd Qu.: 0.100000  
##  Max.   : 1.400000
```

```
head(uschange)
```

```
##         Consumption     Income Production   Savings Unemployment
## 1970 Q1   0.6159862  0.9722610 -2.4527003 4.8103115          0.9
## 1970 Q2   0.4603757  1.1690847 -0.5515251 7.2879923          0.5
## 1970 Q3   0.8767914  1.5532705 -0.3587079 7.2890131          0.5
## 1970 Q4  -0.2742451 -0.2552724 -2.1854549 0.9852296          0.7
## 1971 Q1   1.8973708  1.9871536  1.9097341 3.6577706         -0.1
## 1971 Q2   0.9119929  1.4473342  0.9015358 6.0513418         -0.1
```

```
plot(uschange)
```



Robust parameter estimation, using M estimation

M-estimation defines a weight function such that the estimating equation becomes $\sum_{i=1}^{n} w_i \left(y_i - x'b\right) x_i' = 0$. But the weights depend on the residuals and the residuals on the weights

```
MM=rlm(Consumption ~ Income + Production + Unemployment + Savings,
    data=uschange)
summary(MM)
```

```
##
## Call: rlm(formula = Consumption ~ Income + Production + Unemployment +
##     Savings, data = uschange)
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.86476 -0.15486 -0.01866  0.15042  1.27733
##
## Coefficients:
##              Value    Std. Error t value
## (Intercept)    0.2252   0.0322     6.9877
## Income         0.7525   0.0365    20.5901
## Production     0.0479   0.0224     2.1385
## Unemployment  -0.1556   0.0914    -1.7027
## Savings       -0.0476   0.0024   -19.7573
##
## Residual standard error: 0.2293 on 182 degrees of freedom
```

```
hweights <- data.frame( resid = MM$resid, weight = MM$w)
hweights2 <- hweights[order(MM$w), ]
hweights2[1:15, ]
```

```
##          resid    weight
## 23   1.2773344 0.2414257
## 21   1.0159620 0.3035509
## 52   0.9899880 0.3115075
## 145  0.8928082 0.3454242
## 16  -0.8647649 0.3566168
## 141 -0.8441995 0.3653320
## 63   0.7683315 0.4013531
## 22  -0.7318210 0.4213963
## 154  0.7268712 0.4243003
## 61   0.7200283 0.4282826
## 20  -0.6720621 0.4588468
## 142 -0.6618753 0.4659621
## 129  0.6464282 0.4770988
## 13   0.6172414 0.4996050
## 47  -0.6164972 0.5002263
```

We can see that roughly, as the absolute residual goes down, the weight goes up.