CAPM and Factor Models

Yaping Wang

04/11/2019

Contents

In this Rmd I will revisit the empirical evidence on the CAPM, the Fama-French FF 3-factor model, and the FF 5-factor model. We will use data on US stock market returns. These data are available from Professor Kenneth French's webpage.

```
# Clear everything
rm( list=ls())


# Open the dataset
library(data.table)
ff3<-fread("F-F_Research_Data_Factors.csv", header=TRUE)
ff5<-fread("F-F_Research_Data_5_Factors_2x3.csv", header=TRUE)
pfs<-fread("25_Portfolios_5x5.csv", header=TRUE)

# Converting Percentages to Decimals

cols<-c( "Mkt-RF","SMB","HML","RF")
for (j in cols) set(ff3, j = j, value = ff3[[j]] / 100)

cols2<-c("Mkt-RF","SMB", "HML", "RMW", "CMA", "RF")
for (j in cols2) set(ff5, j = j, value = ff5[[j]] / 100)

cols3<-c( "SMALL LoBM","ME1 BM2", "ME1 BM3",    "ME1 BM4","SMALL HiBM", "ME2 BM1" ,   "ME2 BM2"   ,  "ME
for (j in cols3) set(pfs, j = j, value = pfs[[j]] / 100)
```

1

Calculate for each one of the 25 portfolios its excess return in each month (return minus the risk-free rate) and the sample mean of the time-series of excess returns. Display mean excess returns in a 5x5 Table. Draw a 3-D chart with the values of the 5x5 matrix. What can you say from those results?

mean excess returns table

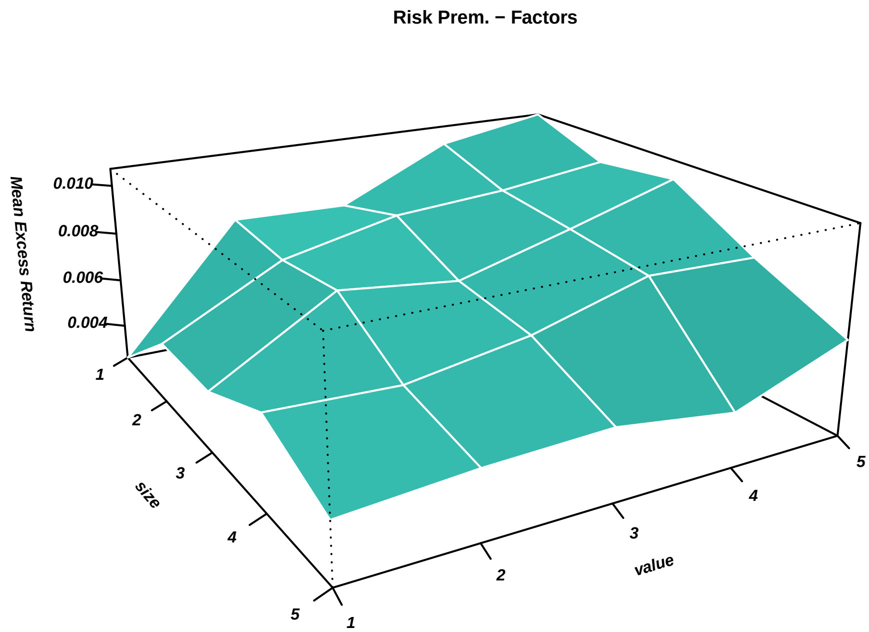
```
# Compute excess return
exR<-pfs[,-c(1)]-ff3[, RF]

# Compute sample mean
exR_mean<-exR[,lapply(.SD, mean)]

# Dispaly in table
mean_matrix<-matrix(exR_mean,nrow = 5, ncol = 5)
colnames(mean_matrix) <- c("ME1","ME2","ME3","ME4","ME5")
rownames(mean_matrix) <- c("BM1","BM2","BM3","BM4","BM5")
mean_matrix
```

```
##     ME1         ME2         ME3         ME4         ME5
## BM1 0.002570662 0.004969903 0.004944982 0.006190803 0.004834288
## BM2 0.007824065 0.00749422  0.007801842 0.006015517 0.005190511
## BM3 0.007726098 0.008562809 0.007278445 0.006796744 0.005352152
## BM4 0.009897741 0.008920026 0.008555186 0.008126045 0.004686283
## BM5 0.01069894  0.009519132 0.009910895 0.007986829 0.006353497
```

```
# Draw the 3D chart
mean_table<-as.matrix.data.frame(t(mean_matrix))
par(ps=6, font.axis=4, font.lab=4)
persp(x=c(1:5), y=c(1:5), mean_table, theta = 60, phi=20, ticktype = "detailed", expand = .4, shade = .
```



What can you say from those results?

Portfolios with small size (markert capitalization) tend to higher excess returns. Similar findings for portfolios with high book-to-market values (BM ratio).

2

Estimated alpha of each portfolio

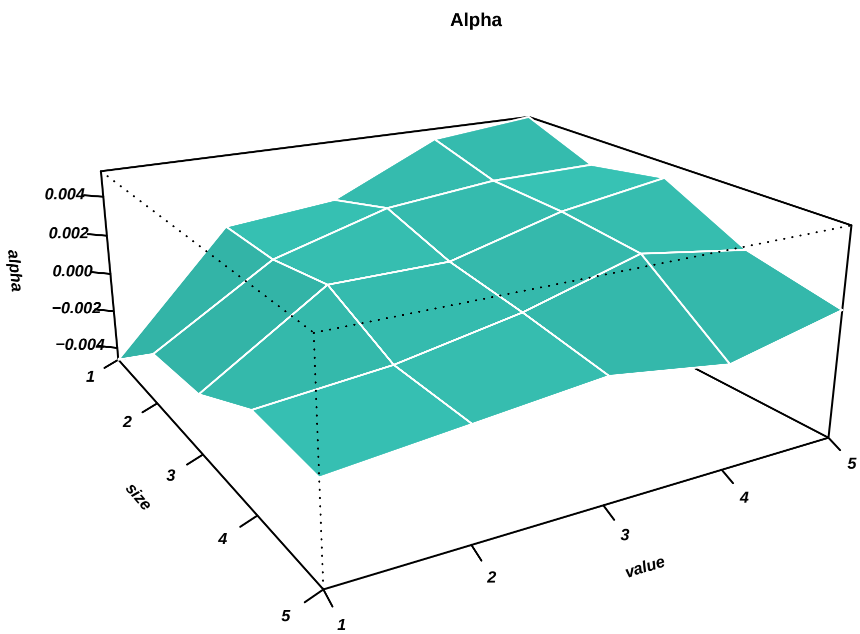
```
# Run a time-series OLS regression for each j
f_aj<-function(x){
  riskprej<-unlist((x-ff3[,"RF"]))
  riskprem<-unlist(ff3[, "Mkt-RF"])
  reg<-lm(riskprej ~ riskprem)
  aj<-summary(reg)$coefficients[1]
  return(aj)
}

# obtain alpha table
alpha<-pfs[,lapply(.SD,f_aj)]
alpha_matrix<-matrix(alpha[,2:26],nrow = 5, ncol = 5)
alpha_matrix
```

```
##        [,1]         [,2]         [,3]          [,4]          [,5]
## [1,] -0.00465166 -0.002122861 -0.001778042 -3.353608e-05 -0.0001369658
```

```
## [2,] 0.001528296 0.001509825  0.002115513  0.0005497773  0.0004482965
## [3,] 0.002092642 0.003179535  0.002157932  0.001663809   0.001039562
## [4,] 0.00470146  0.003763836  0.003664824  0.003287013   0.0002267887
## [5,] 0.005287916 0.003826157  0.004516478  0.00246158    0.001463989
```

```
# 3-D plot
alpha_table<-as.matrix.data.frame(t(alpha_matrix))
par(ps=6, font.axis=4, font.lab=4)
persp(x=c(1:5), y=c(1:5), alpha_table, theta = 60, phi=20, ticktype = "detailed", expand = .4, shade =
```




```
#for (i in c(2:26)){
#  alpha=c()
#  jriskpre<-unlist(pfs[,i]-ff3[,"RF"])
#  mriskpre<-unlist(ff3[, "Mkt-RF"])
#  lm(jriskpre ~ mriskpre)
#  alpha[i]<-summary(lm(jriskpre ~ mriskpre))$coefficients[1]
#}
#alpha

# GRS test
#cov matrix of residuals
f_res<-function(x){
  riskprej<-unlist((x-ff3[,"RF"]))
  riskprem<-unlist(ff3[, "Mkt-RF"])
  reg<-lm(riskprej ~ riskprem)
  res<-summary(reg)$residuals
  return(res)
```

```
}
resi<-pfs[,lapply(.SD,f_res)]
residuals<-resi[,2:26]
covm<-cov(residuals)

#store alphas as a col vector
abreturn<-as.numeric(alpha[,2:26])

#Sharp ratio of the market portfolio
mean_Rm_rf<-mean(ff3$"Mkt-RF")
std_Rm_rf<-sd(ff3$"Mkt-RF")
Msharp<-mean_Rm_rf/std_Rm_rf

f_stat<-(660-25-1)/25*(1/(1+(Msharp)^2))*t(abreturn)%*%solve(covm)%*%(abreturn)
GRS1<-pf(f_stat,25,634,lower.tail=F)
```

What does the table suggest about the validity of the CAPM?

If the CAPM holds, alpha of each portfolio should be equal to zero. However, as the regressions imply, there are significant positive or negtive pricing errors. We can conclude that there are positive alphas, thus it seems that the CAPM cannot fully price our stocks.

What portfolios deviate the most from the CAPM predictions?

The portfolio with the highest value and smallest size (Small HiBM).

Joint Test using GRS t-statistics

The GRS test statistic is 4.576918 and the corresponding $p = 3.943519^{-12} < 1\%$, allowing us to reject the null hypothsis that all alphas are jointly zero at 1% significance level.

3

Test the CAPM: summary of CRS

```
# obtain betas
f_bj<-function(x){
  riskprej<-unlist((x-ff3[,"RF"]))
  riskprem<-unlist(ff3[, "Mkt-RF"])
  reg<-lm(riskprej ~ riskprem)
  bj<-summary(reg)$coefficients[2]
  return(bj)
}

betas<-pfs[,lapply(.SD,f_bj)]
betas1<-betas[,2:26] # drop the month col

# Run a single cross-sectional OLS of mean excess returns on betas
ex<-function(x){
  ex<-unlist((x-ff3[,"RF"]))
  return(ex)
}
exreturn<-pfs[,lapply(.SD,ex)]
```

```r
exR_mean<-exreturn[,lapply(.SD,mean)][,2:26]

y<-unlist(exR_mean)
x_beta<-unlist(betas1)
reg1<-lm(y ~ x_beta)

# Test the CAPM - gamma0
summary(reg1)
```

```
##
## Call:
## lm(formula = y ~ x_beta)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0034011 -0.0011828  0.0005559  0.0012309  0.0035124
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.010816   0.002973   3.637  0.00138 **
## x_beta      -0.003402   0.002719  -1.251  0.22347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.001972 on 23 degrees of freedom
## Multiple R-squared:  0.06372,    Adjusted R-squared:  0.02301
## F-statistic: 1.565 on 1 and 23 DF,  p-value: 0.2235
```

```r
# Test the CAPM - gamma1
mean_Rm_rf<-as.numeric(ff3[,lapply(.SD, mean)][,2])
# alternative: mean_Rm_rf<-mean(ff3$"Mkt-RF")
t_stat1<-(summary(reg1)$coefficients[2]-mean_Rm_rf)/summary(reg1)$coefficients[2,2]
p_gamma1<-pnorm(unlist(t_stat1))
```

γ_0: As is shown above, the estimate of γ_0 is 0.010816 and the corresponding $p = 0.00138 < 1\%$, thus we reject the null hypothesis that $\gamma_0 = 0$.

γ_1: The estimate of γ_1 is -0.003402 and the $\overline{R_m - R_f}$ is 0.005072424. The t statistics is -3.11658 and corresonding $p = 0.0009148101 < 1\%$, thus we reject the null hypothesis that $\gamma_1 = \overline{R_m - R_f}$.

The statistical tests for γ_0 and γ_1 both imply CAPM does not hold.

4

```r
# First step: estimate each stock's beta for each month t
f_bj2<-function(x){
y<-c()
for (i in c(1:624)){
riskprej2<-unlist(x[i:(i+35)]-ff3[(i:(i+35)),"RF"])
riskprem2<-unlist(ff3[(i:(i+35)),"Mkt-RF"])
y[i]<-summary(lm(riskprej2~riskprem2))$coefficients[2]
}
return(y)
}
```

```r
betas2<-pfs[,lapply(.SD,f_bj2)]

betas_Fama<-betas2[,2:26] # drop the month col

# for double check use
#for (i in c(1:624)){
#riskprej2<-unlist(pfs[(i:(i+35)), 3]-ff3[(i:(i+35)),"RF"])
#riskprem2<-unlist(ff3[(i:(i+35)),"Mkt-RF"])
#d[i]<-summary(lm(riskprej2~riskprem2))$coefficients[2]}
#d

# Second step: run cross-sectional regressions for each time t
# compute excess returns at each t for each portfolio
ex<-function(x){
  ex<-unlist((x-ff3[,"RF"]))
  return(ex)
}
exreturn<-pfs[,lapply(.SD,ex)]

fgamma0_t<-c()
fgamma1_t<-c()
for (i in c(1:624)){
  y2<-unlist(exreturn[(36+i),-c(1)])
  x_beta2<-unlist(betas_Fama[i,])
  fgamma0_t[i]<-summary(lm(y2 ~ x_beta2))$coefficients[1]
  fgamma1_t[i]<-summary(lm(y2 ~ x_beta2))$coefficients[2]
}


# Compute Fama-MacBeth gamma0 & gamma1
fgamma0<-mean(fgamma0_t)
fgamma1<-mean(fgamma1_t)

# Compute the Std errors of Fama-MacBeth's estimates
t<-624
fgamma0_t_var<-var(fgamma0_t)
fgamma1_t_var<-var(fgamma1_t)

fgamma0_std<-sqrt((1/t)*fgamma0_t_var)
fgamma1_std<-sqrt((1/t)*fgamma1_t_var)

# Test Fama-MacBeth gamma0 & gamma1
fmean_Rm_rf<-as.numeric(ff3[37:660,lapply(.SD, mean)][,2])
t_stat_f0<-(fgamma0)/fgamma0_std
t_stat_f1<-(fgamma1-fmean_Rm_rf)/fgamma1_std
p_fgamma0<-pnorm(unlist(t_stat_f0),lower.tail=FALSE)
p_fgamma1<-pnorm(unlist(t_stat_f1))
```

The Fama-MacBeth coefficients and their standard errors

$\gamma_0 = 0.008122115$

$\gamma_1 = -0.001276719$

$Std\gamma_0 = 0.002484676$

$Std\gamma_1 = 0.002827005$

Comparing with the OLS ones and Testing the CAPM

Standard errors is a measure of precision of the coefficients we estimated, meaning how much they vary from sample to sample.

Instead of running a single Cross Sectional Regression, running a large number of Cross Sectional Regressions corrects for Cross Sectional correlation and homoskedasticity assumption. It corrects the standard errors which otherwise could have been overrstimated or underestimated due to the fact that a single Cross Sectional Regression overlooks the effects of Cross Sectional Correlation and heteroskedastic errors which are totally possible in real world.

Compared with the results in question 3, the standard error we obtained for γ_0 is lower but it is higher for γ_1. This implies γ_0 tends to vary less within samples while γ_1 tends to vary more compared with γ_0.

The p value for each coefficient implies CAPM does not hold. For γ_0, $p = 0.000539865 < 1\%$ allows us to reject the hypothesis $\gamma_0 = 0$ at 1% significance level which implies there is significant positive alpha. For γ_1, $p = 0.01068794 < 5\%$ allows us to reject the hypothesis $\gamma_1 = \overline{R_m - R_f}$ at 5% significance level.

5

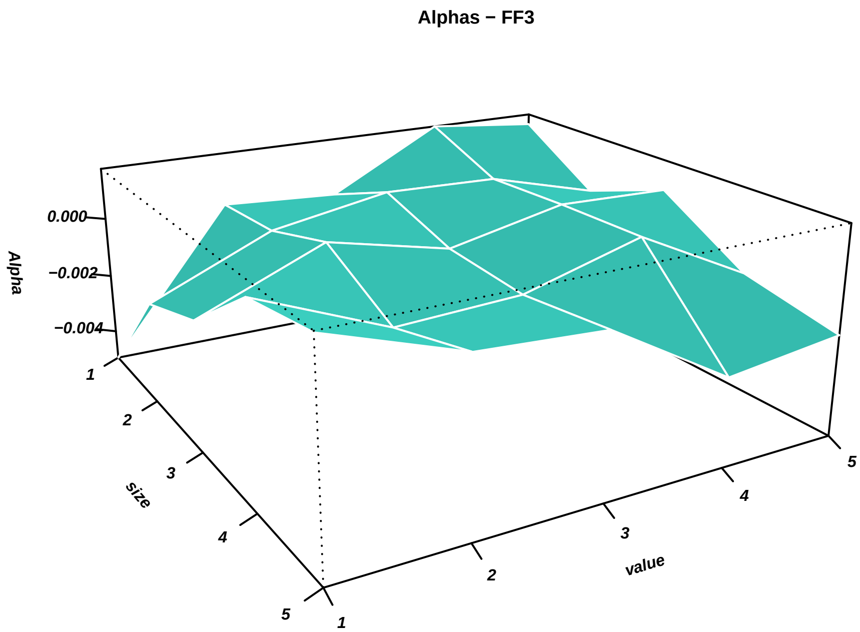
FF3 - estimated alphas and the 3-D plot

```r
# Run a single time-series regression for FF3
f_aj_ff3<-function(x){
  riskprej<-unlist((x-ff3[,"RF"]))
  riskprem<-unlist(ff3[, "Mkt-RF"])
  riskpresmb<-unlist(ff3[, "SMB"])
  riskprehml<-unlist(ff3[, "HML"])
  reg<-lm(riskprej ~ riskprem+riskpresmb+riskprehml)
  aj_ff3<-summary(reg)$coefficients[1]
  return(aj_ff3)
}

# obtain alpha_ff3 table
alpha_ff3<-pfs[,lapply(.SD,f_aj_ff3)]
alpha_ff3_matrix<-matrix(alpha_ff3[,2:26],nrow = 5, ncol = 5)
alpha_ff3_matrix
```

```
##      [,1]          [,2]          [,3]          [,4]          [,5]
## [1,] -0.004975859  -0.001670382  -0.0008288505 0.001198649   0.001666713
## [2,] -0.0001090178 6.279987e-05  0.0008302362  -0.0004439674 0.0003332602
## [3,] -0.0003275903 0.000761636   -4.607832e-05 -0.0002538807 9.084901e-05
## [4,] 0.001707406   0.0006612621  0.000774819   0.000816939   -0.002159445
## [5,] 0.001329362   -0.0003421556 0.0006842319  -0.001047745  -0.001717761
```

```r
# Display alphas of FF3
alpha_ff3_table<-as.matrix.data.frame(t(alpha_ff3_matrix))
par(ps=6, font.axis=4, font.lab=4)
persp(x=c(1:5), y=c(1:5), alpha_ff3_table, theta = 60, phi=20, ticktype = "detailed", expand = .4, shad
```




```
# GRS test for FF3
#cov matrix of residuals
f_res_ff3<-function(x){
  riskprej<-unlist((x-ff3[,"RF"]))
  riskprem<-unlist(ff3[, "Mkt-RF"])
  riskpresmb<-unlist(ff3[, "SMB"])
  riskprehml<-unlist(ff3[, "HML"])
  reg<-lm(riskprej ~ riskprem+riskpresmb+riskprehml)
  res_ff3<-summary(reg)$residuals
  return(res_ff3)
}


resi_ff3<-pfs[,lapply(.SD,f_res_ff3)]
residuals_ff3<-resi_ff3[,2:26]
covm_ff3<-cov(residuals_ff3)

#store alphas as a col vector
abreturn_ff3<-as.numeric(alpha_ff3[,2:26])

#Sharp ratio of the market portfolio
mean_Rm_rf<-mean(ff3$"Mkt-RF")
std_Rm_rf<-sd(ff3$"Mkt-RF")
Msharp<-mean_Rm_rf/std_Rm_rf

f_stat_ff3<-(660-25-1)/25*(1/(1+(Msharp)^2))*t(abreturn_ff3)%*%solve(covm_ff3)%*%(abreturn_ff3)
```

```
p_GRS_FF3<-pf(f_stat_ff3,25,634,lower.tail=F)
```

FF5 - estimated alphas and the 3-D plot

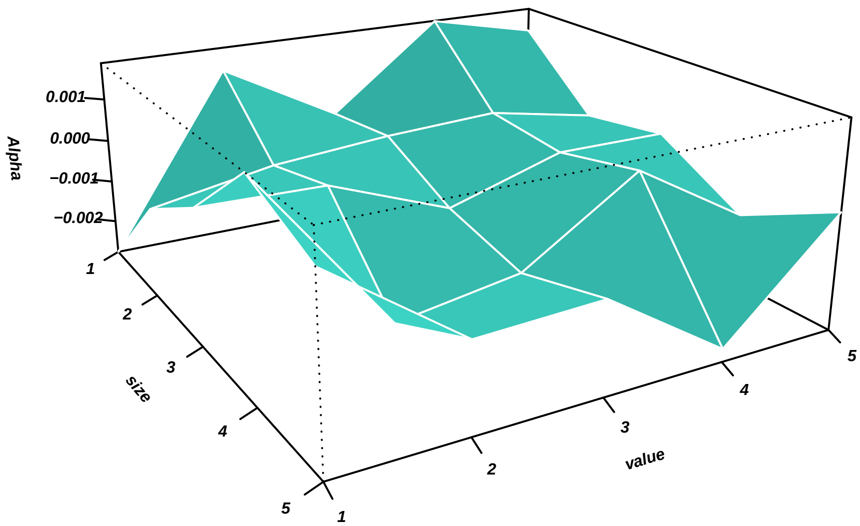
```
# Run a single time-series regression for FF5
f_aj_ff5<-function(x){
  riskprej<-unlist((x-ff5[,"RF"]))
  riskprem<-unlist(ff5[, "Mkt-RF"])
  riskpresmb<-unlist(ff5[, "SMB"])
  riskprehml<-unlist(ff5[, "HML"])
  riskprermw<-unlist(ff5[, "RMW"])
  riskprecma<-unlist(ff5[, "CMA"])
  reg<-lm(riskprej ~ riskprem+riskpresmb+riskprehml+riskprermw+riskprecma)
  aj_ff5<-summary(reg)$coefficients[1]
  return(aj_ff5)
}

# Display alphas of FF5
alpha_ff5<-pfs[,lapply(.SD,f_aj_ff5)]
alpha_ff5_matrix<-matrix(alpha_ff5[,2:26],nrow = 5, ncol = 5)
alpha_ff5_matrix
```

```
##      [,1]          [,2]          [,3]          [,4]         [,5]
## [1,] -0.002787727  -0.0007394001 0.0002413889  0.00185705   0.001154549
## [2,] 0.001301047   -0.0002360439 0.000184862   -0.001747758 -0.0008465007
## [3,] -0.0002080256 -3.497707e-06 -0.0008659603 -0.001321295 -0.000744967
## [4,] 0.001849174   0.0001271777  -5.729468e-05 0.0003761968 -0.002499242
## [5,] 0.001266739   -0.0003722946 -6.901872e-05 -0.001179706 -0.0001488723
```

```
# Display alohas of FF5
alpha_ff5_table<-as.matrix.data.frame(t(alpha_ff5_matrix))
par(ps=6, font.axis=4, font.lab=4)
persp(x=c(1:5), y=c(1:5), alpha_ff5_table, theta = 60, phi=20, ticktype = "detailed", expand = .4, shad
```

Alphas – FF5




```
# GRS test for FF5
#cov matrix of residuals
f_res_ff5<-function(x){
  riskprej<-unlist((x-ff5[,"RF"]))
  riskprem<-unlist(ff5[, "Mkt-RF"])
  riskpresmb<-unlist(ff5[, "SMB"])
  riskprehml<-unlist(ff5[, "HML"])
  riskprermw<-unlist(ff5[, "RMW"])
  riskprecma<-unlist(ff5[, "CMA"])
  reg<-lm(riskprej ~ riskprem+riskpresmb+riskprehml+riskprermw+riskprecma)
  res_ff5<-summary(reg)$residuals
  return(res_ff5)
}


resi_ff5<-pfs[,lapply(.SD,f_res_ff5)]
residuals_ff5<-resi_ff5[,2:26]
covm_ff5<-cov(residuals_ff5)

#store alphas as a col vector
abreturn_ff5<-as.numeric(alpha_ff5[,2:26])

#Sharp ratio of the market portfolio
mean_Rm_rf<-mean(ff5$"Mkt-RF")
std_Rm_rf<-sd(ff5$"Mkt-RF")
Msharp<-mean_Rm_rf/std_Rm_rf
```

```
f_stat_ff5<-(660-25-1)/25*(1/(1+(Msharp)^2))*t(abreturn_ff5)%*%solve(covm_ff5)%*%(abreturn_ff5)
p_GRS_FF5<-pf(f_stat_ff5,25,634,lower.tail=F)
```

Interpretations:

The FF 3 factors and the FF 5 factor models take additional common sources of risks into account, which could possibly decrese pricing errors.

As can be seen from the 3-D graphs, the deviations of pricing errors dramatically decrease compared to the ones we obtained in question 2.

If we check the GRS test statistics for each model, both of them imply there is significant positive alpha. For FF 3 factors model, $t_{GRS} = 3.953142$ and $p = 7.668097^{-10} < 1\%$ allow us to reject the hypothesis $\alpha = 0$ at 1% significance level. For FF 5 factors model, $t_{GRS} = 3.427978$ and $p = 5.902528e^{-8} < 1\%$ also allow us to reject the hypothesis $\alpha = 0$ at 1% significance level.