**Exploration**

---

title: "Exploration"

author: "Panther"

date: "Jan 28,2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);

# here data read is being processed

t <- read\_csv("../data/data.csv")

```

#Process Data

```{r Process Data}

GSPC <- read\_csv("../data/^GSPC.csv")

LIBOR <- read\_csv("../data/USD3MTD156N.csv")

Treasury <- read\_csv("../data/DTB3.csv")

VIX <- read\_csv("../data/^VIX.csv")

t<-cbind(GSPC,Treasury$DTB3[match(GSPC$Date,Treasury$DATE)])

t<-cbind(t,VIX$Close[match(t$Date,VIX$Date)])

t<-t[,c(1,5,8,9)]

t<- t[-1:(-10054+120),] # extra 120 for training signal, will remove later

colnames(t)[3] <- "Free\_rate"

colnames(t)[4] <- "VIX"

t$Free\_rate = as.numeric(as.character(t$Free\_rate))

```

#Missing Data

```{r Missing Data}

sum(is.na(t))

sum(is.na(t$Close))

sum(is.na(t$Free\_rate))

sum(is.na(t$VIX))

which(is.na(t$Free\_rate))

#count(which(is.na(t$Free\_rate)))

temp <- which(is.na(t$Free\_rate))

for (i in temp)

t$Free\_rate[i] <- mean(t$Free\_rate[(i-5):(i+5)], na.rm=TRUE)

sum(is.na(t$Free\_rate))

```

#Moving Average and momentum signal

```{r momentum signal}

t$SMA60 <- SMA(t$Close,60)

t$SMA120 <- SMA(t$Close,120)

t<- t[-1:-120,] # here remove training 120

for (i in seq\_along(t$SMA120)){

if (t$SMA60[i]>t$SMA120[i])

t$signal[i] = 1 # signal = 1 is buy singal

else if (t$SMA60[i]<t$SMA120[i])

t$signal[i] = 0

else

t$signal[i] = t$singal[i-1]

}

```

#Plot SPC and moving average

```{r Plot index and moving average}

t$POSIXct\_Date <- as.POSIXct(t$Date)

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for all data")

a = 4500

c = 7069

ggplot(t[a:c,],aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for 10 years")

```

Momentum effect

Strong autocorrelation

```{r Momentum effect}

acf(t$Close,lag.max = 100,type='correlation')

pacf(t$Close,lag.max = 20)

```

## Create portfolios

Absolute Return = shift between one lag of close prices

Percentage Return = Absolute Return / last day close price

??Relative Return = Percentage return \* anchor price(first day price of index)

??Better not use due to different products and performences

#portfolio 1

Buy/sell 1 index on today's closing and close position next day

```{r portfolio 1}

t$port1\_absreturn[1] = NA

t$port1\_Perreturn[1] = NA

#t$port1\_Relreturn[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

t$port1\_absreturn[i+1] = t$Close[i+1]-t$Close[i]

t$port1\_Perreturn[i+1] = (t$Close[i+1]-t$Close[i]) / t$Close[i]

# t$port1\_Relreturn[i+1] = t$port1\_Perreturn[i+1] \* t$Close[1]

}

else{

t$port1\_absreturn[i+1] = t$Close[i]-t$Close[i+1]

t$port1\_Perreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i]

# t$port1\_Relreturn[i+1] = t$port1\_Perreturn[i+1] \* t$Close[1]

}

}

```

#portfolio 2

Buy/sell at the money option(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

1/100 option means underline asset is 1 index

```{r portfolio 2}

t$port2\_absreturn[1] = NA

t$port2\_Perreturn[1] = NA

#t$port2\_Relreturn[1] = NA

#if (t$signal[1] == 1) {

# cc <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

# } else {

# cc <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))}

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

aa <- GBSOption("c",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

bb <- GBSOption("c",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

} else{

aa <- GBSOption("p",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

bb <- GBSOption("p",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

}

t$port2\_absreturn[i+1] = aa@price-bb@price

t$port2\_Perreturn[i+1] = (aa@price-bb@price) / bb@price

# t$port2\_Relreturn[i+1] = t$port2\_Perreturn[i+1] \* t$Close[1]

}

```

#portfolio 3

Buy at the money straddle(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

1/100 option means underline asset is 1 index

```{r portfolio 3}

t$port3\_absreturn[1] = NA

t$port3\_Perreturn[1] = NA

#t$port3\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

aa1 <- GBSOption("c",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

aa2 <- GBSOption("p",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

bb1 <- GBSOption("c",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

bb2 <- GBSOption("p",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

t$port3\_absreturn[i+1] = (aa1@price+aa2@price-bb1@price-bb2@price)

t$port3\_Perreturn[i+1] = (aa1@price+aa2@price-bb1@price-bb2@price) / (bb1@price+bb2@price)

# t$port3\_Relreturn[i+1] = t$port3\_Perreturn[i+1] \* t$Close[1]

}

```

#portfolio 4

Buy at the money straddle(90 days maturity but remain life decrease until rebalance at maturity)

- on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

1/100 option means underline asset is 1 index

```{r portfolio 4}

t$port4\_absreturn[1] = NA

t$port4\_Perreturn[1] = NA

#t$port4\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

#date.lookup <- format(seq(as.Date("2000-01-02"), as.Date("2018-1-19"), by = "1 day"))

#date.except <- subset(b, !(y %in% a$x))

#match("2024-01-19", date.lookup)

count=0 #count days passed

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (as.numeric(difftime(t$Date[i+1],t$Date[i])) == 1){ #workday

count = count+1}

else { # not workday

count = count+as.numeric(difftime(t$Date[i+1],t$Date[i])) }

if (count <= 89){

Remaina = 90 - count # today's remain

Remainb = Remaina + 1} # previous day's remain

else if (count >= 90) { # expired already, adjuste to 90 days passed

Remaina = 0 # today's remain

Remainb = 1 # previous day's remain

count = 0 } # after expired, rebalance

aa1 <- GBSOption("c",t$Close[i+1],t$Close[i],0.246575342\*Remaina/90,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

aa2 <- GBSOption("p",t$Close[i+1],t$Close[i],0.246575342\*Remaina/90,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

bb1 <- GBSOption("c",t$Close[i],t$Close[i],0.246575342\*Remainb/90,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

bb2 <- GBSOption("p",t$Close[i],t$Close[i],0.246575342\*Remainb/90,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

t$port4\_absreturn[i+1] = (aa1@price+aa2@price-bb1@price-bb2@price)

t$port4\_Perreturn[i+1] = (aa1@price+aa2@price-bb1@price-bb2@price) / (bb1@price+bb2@price)

# t$port4\_Relreturn[i+1] = t$port4\_Perreturn[i+1] \* t$Close[1]

}

```

#Dump first day NA return

```{r}

t <- t[-1,]

```

#Line Plot Absolute Return

```{r Line Plot Absolute Return}

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_absreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_absreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_absreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_absreturn))

```

#Line Plot Percentage Return

```{r Line Plot Percentage Return}

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_Perreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_Perreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_Perreturn))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_Perreturn))

```

# Scatter of Absolute Return

```{r Scatter Plot Absolute Return}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_absreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_absreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_absreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port4\_absreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_absreturn,color="Port1"),alpha=0.2) +

geom\_point(aes(y=port2\_absreturn,color="Port2"),alpha=0.2) +

geom\_point(aes(y=port3\_absreturn,color="Port3"),alpha=0.2) +

geom\_point(aes(y=port4\_absreturn,color="Port4"),alpha=0.2) +

theme\_bw() +

ggtitle("absreturn Return comparing for four portfolios")

a = 4500

c = 7068

ggplot(t[a:c,],aes(x=Date)) +

geom\_point(aes(y=port1\_absreturn,color="Port1"),alpha=0.5,shape= 1) +

geom\_point(aes(y=port2\_absreturn,color="Port2"),alpha=0.5,shape= 0) +

geom\_point(aes(y=port3\_absreturn,color="Port3"),alpha=0.5,shape= 2) +

geom\_point(aes(y=port4\_absreturn,color="Port4"),alpha=0.5,shape= 5) +

theme\_bw() +

ggtitle("absreturn Return comparing for four portfolios in 10 years")

```

# Scatter of Percentage Return

```{r Scatter Plot Percentage Return}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port4\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Perreturn,color="Port1"),alpha=0.2) +

geom\_point(aes(y=port2\_Perreturn,color="Port2"),alpha=0.2) +

geom\_point(aes(y=port3\_Perreturn,color="Port3"),alpha=0.2) +

geom\_point(aes(y=port4\_Perreturn,color="Port4"),alpha=0.2) +

theme\_bw() +

ggtitle("percentage Return comparing for four portfolios")

a = 4500

c = 7068

ggplot(t[a:c,],aes(x=Date)) +

geom\_point(aes(y=port1\_Perreturn,color="Port1"),alpha=0.5,shape= 1) +

geom\_point(aes(y=port2\_Perreturn,color="Port2"),alpha=0.5,shape= 0) +

geom\_point(aes(y=port3\_Perreturn,color="Port3"),alpha=0.5,shape= 2) +

geom\_point(aes(y=port4\_Perreturn,color="Port4"),alpha=0.5,shape= 5) +

theme\_bw() +

ggtitle("percentage Return comparing for four portfolios in 10 years")

```

#Histogram of Absolute Return

```{r Histogram Plot Absolute Return}

ggplot(t,aes(x = port1\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_absreturn),sd = sd(t$port1\_absreturn))) +

theme\_bw()

ggplot(t,aes(x = port2\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_absreturn),sd = sd(t$port2\_absreturn))) +

theme\_bw()

ggplot(t,aes(x = port3\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_absreturn),sd = sd(t$port3\_absreturn))) +

theme\_bw()

ggplot(t,aes(x = port4\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_absreturn),sd = sd(t$port4\_absreturn))) +

theme\_bw()

a = 4500

c = 7068

print("Absolute Return in 10 years")

ggplot(t[a:c,],aes(x = port1\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_absreturn),sd = sd(t$port1\_absreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port2\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_absreturn),sd = sd(t$port2\_absreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port3\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_absreturn),sd = sd(t$port3\_absreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port4\_absreturn)) +

geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_absreturn),sd = sd(t$port4\_absreturn))) +

theme\_bw()

```

#Histogram of Percentage Return

```{r Histogram Plot Percentage Return}

ggplot(t,aes(x = port1\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Perreturn),sd = sd(t$port1\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port2\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Perreturn),sd = sd(t$port2\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port3\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Perreturn),sd = sd(t$port3\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port4\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Perreturn),sd = sd(t$port4\_Perreturn))) +

theme\_bw()

a = 4500

c = 7068

print("Percentage Return in 10 years")

ggplot(t[a:c,],aes(x = port1\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Perreturn),sd = sd(t$port1\_Perreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port2\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Perreturn),sd = sd(t$port2\_Perreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port3\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Perreturn),sd = sd(t$port3\_Perreturn))) +

theme\_bw()

ggplot(t[a:c,],aes(x = port4\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Perreturn),sd = sd(t$port4\_Perreturn))) +

theme\_bw()

```

#QQ plot of Absolute Return

```{r QQ Plot Absolute Return}

ggplot(t, aes(sample=port1\_absreturn)) + stat\_qq()

ggplot(t, aes(sample=port2\_absreturn)) + stat\_qq()

ggplot(t, aes(sample=port3\_absreturn)) + stat\_qq()

ggplot(t, aes(sample=port4\_absreturn)) + stat\_qq()

a = 4500

c = 7068

print("Absolute Return in 10 years")

ggplot(t[a:c,], aes(sample=port1\_absreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port2\_absreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port3\_absreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port4\_absreturn)) + stat\_qq()

```

#QQ plot of Percentage Return

```{r QQ Plot Percentage Return}

ggplot(t, aes(sample=port1\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port2\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port3\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port4\_Perreturn)) + stat\_qq()

a = 4500

c = 7068

print("Percentage Return in 10 years")

ggplot(t[a:c,], aes(sample=port1\_Perreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port2\_Perreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port3\_Perreturn)) + stat\_qq()

ggplot(t[a:c,], aes(sample=port4\_Perreturn)) + stat\_qq()

```

#Statistics of Absolute Return

```{r Statistics of Absolute Return}

cat(red("absreturn Statistics in all time\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_absreturn)

describe(t$port1\_absreturn)

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_absreturn)

describe(t$port2\_absreturn)

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_absreturn)

describe(t$port3\_absreturn)

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_absreturn)

describe(t$port4\_absreturn)

a = 4500

c = 7068

cat(red("absreturn Statistics in 10 years\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_absreturn[a:c])

describe(t$port1\_absreturn[a:c])

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_absreturn[a:c])

describe(t$port2\_absreturn[a:c])

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_absreturn[a:c])

describe(t$port3\_absreturn[a:c])

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_absreturn[a:c])

describe(t$port4\_absreturn[a:c])

```

#Statistics of Percentage Return

```{r Statistics of Percentage Return}

cat(red("Percentage Statistics in all time\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Perreturn)

describe(t$port1\_Perreturn)

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Perreturn)

describe(t$port2\_Perreturn)

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Perreturn)

describe(t$port3\_Perreturn)

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Perreturn)

describe(t$port4\_Perreturn)

a = 4500

c = 7068

cat(red("Percentage Statistics in 10 years\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Perreturn[a:c])

describe(t$port1\_Perreturn[a:c])

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Perreturn[a:c])

describe(t$port2\_Perreturn[a:c])

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Perreturn[a:c])

describe(t$port3\_Perreturn[a:c])

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Perreturn[a:c])

describe(t$port4\_Perreturn[a:c])

```

#Statistics Comparison

```{r Statistics table}

base <- t[,c(9,11,13,15,10,12,14,16)]

colnames(base) <- c("port1\_absreturn","port2\_absreturn","port3\_absreturn","port4\_absreturn","port1\_Perreturn","port2\_Perreturn","port3\_Perreturn","port4\_Perreturn")

#base <- t[,c(9,11,13,15)]

#colnames(base) <- c("port1\_absreturn","port2\_absreturn","port3\_absreturn","port4\_absreturn")

stat <- sapply(base,sum) %>%

rbind(sapply(base,mean)) %>%

rbind(sapply(base,mean)\*252) %>%

rbind(sapply(base,var)) %>%

rbind(sapply(base,skew)) %>%

rbind(sapply(base,kurtosis))

rownames(stat) <- c("sum","mean","annual mean","var","skew","kurtosis")

print("Statistics in all time")

print(stat)

a = 4500

c = 7068

stat2 <- sapply(base[a:c,],sum) %>%

rbind(sapply(base[a:c,],mean)) %>%

rbind(sapply(base[a:c,],mean)\*252) %>%

rbind(sapply(base[a:c,],var)) %>%

rbind(sapply(base[a:c,],skew)) %>%

rbind(sapply(base[a:c,],kurtosis))

rownames(stat2) <- c("sum","mean","annual mean","var","skew","kurtosis")

print("Statistics in 10 years")

print(stat2)

```

#Culmulative Return

See how the cost of options and straddles ruined your return

```{r Culmulative returns}

t$port1\_cum\_abs\_return<-cumsum(t$port1\_absreturn)

t$port2\_cum\_abs\_return<-cumsum(t$port2\_absreturn)

t$port3\_cum\_abs\_return<-cumsum(t$port3\_absreturn)

t$port4\_cum\_abs\_return<-cumsum(t$port4\_absreturn)

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_abs\_return,color="Port1")) +

geom\_line(aes(y=port2\_cum\_abs\_return,color="Port2")) +

geom\_line(aes(y=port3\_cum\_abs\_return,color="Port3")) +

geom\_line(aes(y=port4\_cum\_abs\_return,color="Port4")) +

theme\_bw() +

ggtitle("abs Return comparing in all time")

a = 4500

c = 7068

t$port1\_cum\_abs\_return2[a:c]<-cumsum(t$port1\_absreturn[a:c])

t$port2\_cum\_abs\_return2[a:c]<-cumsum(t$port2\_absreturn[a:c])

t$port3\_cum\_abs\_return2[a:c]<-cumsum(t$port3\_absreturn[a:c])

t$port4\_cum\_abs\_return2[a:c]<-cumsum(t$port4\_absreturn[a:c])

ggplot(t[a:c,],aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_abs\_return2,color="Port1")) +

geom\_line(aes(y=port2\_cum\_abs\_return2,color="Port2")) +

geom\_line(aes(y=port3\_cum\_abs\_return2,color="Port3")) +

geom\_line(aes(y=port4\_cum\_abs\_return2,color="Port4")) +

theme\_bw() +

ggtitle("abs Return comparing in 10 years")

```

#Realized Volatility on abs return

Realized Vol is limited, 10 years has upper boundary

```{r Realized Volatility}

t$port1\_real\_vol <- 252\*cumsum(t$port1\_absreturn^2)/seq(length(t$port1\_absreturn)) %>% sqrt()

t$port2\_real\_vol <- 252\*cumsum(t$port2\_absreturn^2)/seq(length(t$port2\_absreturn)) %>% sqrt()

t$port3\_real\_vol <- 252\*cumsum(t$port3\_absreturn^2)/seq(length(t$port3\_absreturn)) %>% sqrt()

t$port4\_real\_vol <- 252\*cumsum(t$port4\_absreturn^2)/seq(length(t$port4\_absreturn)) %>% sqrt()

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_real\_vol,color="Port1")) +

geom\_line(aes(y=port2\_real\_vol,color="Port2")) +

geom\_line(aes(y=port3\_real\_vol,color="Port3")) +

geom\_line(aes(y=port4\_real\_vol,color="Port4")) +

theme\_bw() +

ggtitle("Realized Volatility in all time")

a = 4500

c = 7068

t$port1\_real\_vol2[a:c] <- 252\*cumsum(t$port1\_absreturn[a:c]^2)/seq(length(t$port1\_absreturn[a:c])) %>% sqrt()

t$port2\_real\_vol2[a:c] <- 252\*cumsum(t$port2\_absreturn[a:c]^2)/seq(length(t$port2\_absreturn[a:c])) %>% sqrt()

t$port3\_real\_vol2[a:c] <- 252\*cumsum(t$port3\_absreturn[a:c]^2)/seq(length(t$port3\_absreturn[a:c])) %>% sqrt()

t$port4\_real\_vol2[a:c] <- 252\*cumsum(t$port4\_absreturn[a:c]^2)/seq(length(t$port4\_absreturn[a:c])) %>% sqrt()

ggplot(t[a:c,],aes(x=Date)) +

geom\_line(aes(y=port1\_real\_vol2,color="Port1")) +

geom\_line(aes(y=port2\_real\_vol2,color="Port2")) +

geom\_line(aes(y=port3\_real\_vol2,color="Port3")) +

geom\_line(aes(y=port4\_real\_vol2,color="Port4")) +

theme\_bw() +

ggtitle("Realized Volatility in 10 years")

```

#Rolling Analysis annually(252 days)

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis annually}

port1 <- xts(x=t$port1\_absreturn, order.by=t$Date)

port2 <- xts(x=t$port2\_absreturn, order.by=t$Date)

port3 <- xts(x=t$port3\_absreturn, order.by=t$Date)

port4 <- xts(x=t$port4\_absreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,252,gap=252,by=1,FUN="mean")

t$port2\_roll\_mean <- apply.rolling(port2,252,gap=252,by=1,FUN="mean")

t$port3\_roll\_mean <- apply.rolling(port3,252,gap=252,by=1,FUN="mean")

t$port4\_roll\_mean <- apply.rolling(port4,252,gap=252,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,252,gap=252,by=1,FUN="sd")

t$port2\_roll\_vol <- apply.rolling(port2,252,gap=252,by=1,FUN="sd")

t$port3\_roll\_vol <- apply.rolling(port3,252,gap=252,by=1,FUN="sd")

t$port4\_roll\_vol <- apply.rolling(port4,252,gap=252,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,252)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,252)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,252)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,252)

```

#Rolling Analysis 60 days

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis 60 days}

port1 <- xts(x=t$port1\_absreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,60,gap=60,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,60,gap=60,by=1,FUN="sd")

port2 <- xts(x=t$port2\_absreturn, order.by=t$Date)

t$port2\_roll\_mean <- apply.rolling(port2,60,gap=60,by=1,FUN="mean")

t$port2\_roll\_vol <- apply.rolling(port2,60,gap=60,by=1,FUN="sd")

port3 <- xts(x=t$port3\_absreturn, order.by=t$Date)

t$port3\_roll\_mean <- apply.rolling(port3,60,gap=60,by=1,FUN="mean")

t$port3\_roll\_vol <- apply.rolling(port3,60,gap=60,by=1,FUN="sd")

port4 <- xts(x=t$port4\_absreturn, order.by=t$Date)

t$port4\_roll\_mean <- apply.rolling(port4,60,gap=60,by=1,FUN="mean")

t$port4\_roll\_vol <- apply.rolling(port4,60,gap=60,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,60)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,60)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,60)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,60)

```

**Exploration 2**

---

title: "Exploration"

author: "Panther"

date: "Jan 28,2018"

output:

html\_document: default

pdf\_document: default

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);

# here data read is being processed

#t <- read\_csv("../data/data\_div.csv")

#t <- read\_csv("../Data/data.csv")

t <- read\_csv("F:/Dropbox/[IAQF]/Data/data.csv")

# t$Date <- strptime(as.character(t$Date), "%m/%d/%Y")

# format(t$Date, "%Y-%m-%d")

# t$Date <- as.Date(t$Date)

```

#Process Data

```{r Process Data}

GSPC <- read\_csv("../data/^GSPC.csv")

LIBOR <- read\_csv("../data/USD3MTD156N.csv")

Treasury <- read\_csv("../data/DTB3.csv")

VIX <- read\_csv("../data/^VIX.csv")

#t$newDate <- strptime(as.character(t$Date), "%m/%d/%Y")

#format(t$newDate, "%Y-%m-%d")

#t$Date <- as.Date(t$newDate)

#t <- t[-10]

t<-cbind(GSPC,Treasury$DTB3[match(GSPC$Date,Treasury$DATE)])

t<-cbind(t,VIX$Close[match(t$Date,VIX$Date)])

t<-t[,c(1,5,8,9)]

t<- t[-1:(-10054+120),] # extra 120 for training signal, will remove later

colnames(t)[3] <- "Free\_rate"

colnames(t)[4] <- "VIX"

t$Free\_rate = as.numeric(as.character(t$Free\_rate))

```

#Missing Data

```{r Missing Data}

sum(is.na(t))

sum(is.na(t$Close))

sum(is.na(t$Free\_rate))

sum(is.na(t$VIX))

which(is.na(t$Free\_rate))

#count(which(is.na(t$Free\_rate)))

temp <- which(is.na(t$Free\_rate))

for (i in temp)

t$Free\_rate[i] <- mean(t$Free\_rate[(i-5):(i+5)], na.rm=TRUE)

sum(is.na(t$Free\_rate))

```

#Moving Average and momentum signal

```{r momentum signal}

t$SMA60 <- SMA(t$Close,60)

t$SMA120 <- SMA(t$Close,120)

t<- t[-1:-120,] # here remove training 120

for (i in seq\_along(t$SMA120)){

if (t$SMA60[i]>t$SMA120[i])

t$signal[i] = 1 # signal = 1 is buy singal

else if (t$SMA60[i]<t$SMA120[i])

t$signal[i] = 0

else

t$signal[i] = t$singal[i-1]

}

```

#Plot SPC and moving average

```{r Plot index and moving average}

t$POSIXct\_Date <- as.POSIXct(t$Date)

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for all data")

a = 4500

c = 7069

ggplot(t[a:c,],aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for 10 years")

```

## Create portfolios

Absolute Return = shift between one lag of close prices

Percentage Return = Absolute Return / last day close price

Relative Return = Percentage return \* initiate price($100,000)

We use Relative Return as analysis object

#portfolio 1

Buy/short 1 index on today's closing and close position next day

Assume when you short today, you buy an index for paying back next day

Recieve dividend while buying

```{r portfolio 1}

#t$port1\_absreturn[1] = NA

#t$port1\_Perreturn[1] = NA

t$port1\_Relreturn[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

# t$port1\_absreturn[i+1] = t$Close[i+1]-t$Close[i]

# t$port1\_Perreturn[i+1] = (t$Close[i+1]-t$Close[i]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i+1]-t$Close[i]+t$Close[i]\*t$DivYield[i]/365\*as.numeric(difftime(t$Date[i+1],t$Date[i]))) / t$Close[i] \* 100000

}

else{

# t$port1\_absreturn[i+1] = t$Close[i]-t$Close[i+1]

# t$port1\_Perreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i] \* 100000

}

}

```

#portfolio 2

Buy at the money call/put option(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 2}

#t$port2\_absreturn[1] = NA

#t$port2\_Perreturn[1] = NA

t$port2\_Relreturn[1] = NA

#if (t$signal[1] == 1) {

# cc <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

# } else {

# cc <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

# for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

#

# if (t$signal[i]== 1){

# aa <- GBSOption("c",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

# bb <- GBSOption("c",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

# } else{

# aa <- GBSOption("p",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

# bb <- GBSOption("p",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

# }

# t$port2\_absreturn[i+1] = aa@price-bb@price

# t$port2\_Perreturn[i+1] = (aa@price-bb@price) / bb@price

# # t$port2\_Relreturn[i+1] = t$port2\_Perreturn[i+1] \* t$Close[1]

# }

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

aa <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("call", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

} else {

aa <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("put", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

}

# t$port2\_absreturn[i+1] = aa$value-bb$value

# t$port2\_Perreturn[i+1] = (aa$value-bb$value) / bb$value

t$port2\_Relreturn[i+1] = (aa$value-bb$value) / bb$value \* 100000

}

```

#portfolio 3

Buy at the money straddle(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 3}

#t$port3\_absreturn[1] = NA

#t$port3\_Perreturn[1] = NA

t$port3\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

aa1 <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

aa2 <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb1 <- EuropeanOption("call", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

bb2 <- EuropeanOption("put", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

# t$port3\_absreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value)

# t$port3\_Perreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)

t$port3\_Relreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)\*100000

}

```

#portfolio 4

Buy at the money straddle(90 days maturity but remain life decrease until rebalance at maturity)

- on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 4}

#t$port4\_absreturn[1] = NA

#t$port4\_Perreturn[1] = NA

t$port4\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

#date.lookup <- format(seq(as.Date("2000-01-02"), as.Date("2018-1-19"), by = "1 day"))

#date.except <- subset(b, !(y %in% a$x))

#match("2024-01-19", date.lookup)

count=0 #count days passed

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (as.numeric(difftime(t$Date[i+1],t$Date[i])) == 1){ #workday

count = count+1}

else { # not workday

count = count+as.numeric(difftime(t$Date[i+1],t$Date[i])) }

if (count <= 89){

Remaina = 90 - count # today's remain

Remainb = Remaina + 1} # previous day's remain

else if (count >= 90) { # expired already, adjuste to 90 days passed

Remaina = 0 # today's remain

Remainb = 1 # previous day's remain

count = 0 } # after expired, rebalance

aa1 <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), 0.246575342\*Remaina/90, (t$VIX[i+1]/100),0, 0)

aa2 <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$DivYield[i+1], (t$Free\_rate[i+1]/100), 0.246575342\*Remaina/90, (t$VIX[i+1]/100),0, 0)

bb1 <- EuropeanOption("call", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342\*Remainb/90, (t$VIX[i]/100),0, 0)

bb2 <- EuropeanOption("put", t$Close[i], t$Close[i],t$DivYield[i], (t$Free\_rate[i]/100), 0.246575342\*Remainb/90, (t$VIX[i]/100),0, 0)

# t$port4\_absreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value)

# t$port4\_Perreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)

t$port4\_Relreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)\*100000

}

```

#Dump first day NA return

```{r}

t <- t[-1,]

```

#Momentum effect of portfolio returns

Obvious autocorrelation, large ARMA parameters

```{r Momentum effect}

acf(t$port1\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port1\_Relreturn,lag.max = 50)

acf(t$port2\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port2\_Relreturn,lag.max = 50)

acf(t$port3\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port3\_Relreturn,lag.max = 50)

acf(t$port4\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port4\_Relreturn,lag.max = 50)

```

#Line Plot Absolute Return

```{r Line Plot Absolute Return}

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port1\_absreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port2\_absreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port3\_absreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port4\_absreturn))

```

#Line Plot Percentage Return

```{r Line Plot Percentage Return}

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port1\_Perreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port2\_Perreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port3\_Perreturn))

#ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=port4\_Perreturn))

```

# Scatter of RelReturn

```{r Scatter Plot RelReturn}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port4\_Relreturn),size =1, shape= 1)

# ylim(-250000,350000)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Relreturn,color="Port1"),alpha=0.2) +

geom\_point(aes(y=port2\_Relreturn,color="Port2"),alpha=0.2) +

geom\_point(aes(y=port3\_Relreturn,color="Port3"),alpha=0.2) +

geom\_point(aes(y=port4\_Relreturn,color="Port4"),alpha=0.2) +

ylim(-100000,200000) +

theme\_bw() +

ggtitle("Relreturn comparing for four portfolios")

#a = 4500

#c = 7068

#ggplot(t[a:c,],aes(x=Date)) +

# geom\_point(aes(y=port1\_absreturn,color="Port1"),alpha=0.5,shape= 1) +

# geom\_point(aes(y=port2\_absreturn,color="Port2"),alpha=0.5,shape= 0) +

# geom\_point(aes(y=port3\_absreturn,color="Port3"),alpha=0.5,shape= 2) +

# geom\_point(aes(y=port4\_absreturn,color="Port4"),alpha=0.5,shape= 5) +

# theme\_bw() +

# ggtitle("absreturn Return comparing for four portfolios in 10 years")

```

# Scatter of Percentage Return

```{r Scatter Plot Percentage Return}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port4\_Perreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Perreturn,color="Port1"),alpha=0.2) +

geom\_point(aes(y=port2\_Perreturn,color="Port2"),alpha=0.2) +

geom\_point(aes(y=port3\_Perreturn,color="Port3"),alpha=0.2) +

geom\_point(aes(y=port4\_Perreturn,color="Port4"),alpha=0.2) +

theme\_bw() +

ggtitle("percentage Return comparing for four portfolios")

#a = 4500

#c = 7068

#ggplot(t[a:c,],aes(x=Date)) +

# geom\_point(aes(y=port1\_Perreturn,color="Port1"),alpha=0.5,shape= 1) +

# geom\_point(aes(y=port2\_Perreturn,color="Port2"),alpha=0.5,shape= 0) +

# geom\_point(aes(y=port3\_Perreturn,color="Port3"),alpha=0.5,shape= 2) +

# geom\_point(aes(y=port4\_Perreturn,color="Port4"),alpha=0.5,shape= 5) +

# theme\_bw() +

# ggtitle("percentage Return comparing for four portfolios in 10 years")

```

#Histogram of RelReturn

Portfolio 2 is strongly non-central and skewed

Portfolio 4 is pretty discrete

```{r Histogram Plot RelReturn}

ggplot(t,aes(x = port1\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Relreturn),sd = sd(t$port1\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port2\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Relreturn),sd = sd(t$port2\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port3\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Relreturn),sd = sd(t$port3\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port4\_Relreturn)) +

geom\_histogram(aes(binwidth = 1,y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Relreturn),sd = sd(t$port4\_Relreturn))) +

# xlim(-10000,10000) +

theme\_bw()

#a = 4500

#c = 7068

#print("Absolute Return in 10 years")

#ggplot(t[a:c,],aes(x = port1\_absreturn)) +

# geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_absreturn),sd = sd(t$port1\_absreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port2\_absreturn)) +

# geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_absreturn),sd = sd(t$port2\_absreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port3\_absreturn)) +

# geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_absreturn),sd = sd(t$port3\_absreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port4\_absreturn)) +

# geom\_histogram(binwidth = 1,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_absreturn),sd = sd(t$port4\_absreturn))) +

# theme\_bw()

```

#Histogram of Percentage Return

```{r Histogram Plot Percentage Return}

ggplot(t,aes(x = port1\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Perreturn),sd = sd(t$port1\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port2\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Perreturn),sd = sd(t$port2\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port3\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Perreturn),sd = sd(t$port3\_Perreturn))) +

theme\_bw()

ggplot(t,aes(x = port4\_Perreturn)) +

geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Perreturn),sd = sd(t$port4\_Perreturn))) +

theme\_bw()

#a = 4500

#c = 7068

#print("Percentage Return in 10 years")

#ggplot(t[a:c,],aes(x = port1\_Perreturn)) +

# geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Perreturn),sd = sd(t$port1\_Perreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port2\_Perreturn)) +

# geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Perreturn),sd = sd(t$port2\_Perreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port3\_Perreturn)) +

# geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Perreturn),sd = sd(t$port3\_Perreturn))) +

# theme\_bw()

#ggplot(t[a:c,],aes(x = port4\_Perreturn)) +

# geom\_histogram(binwidth = 0.001,aes(y=..density..,fill=..count..)) +

# stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Perreturn),sd = sd(t$port4\_Perreturn))) +

# theme\_bw()

```

#QQ plot of Relreturn

```{r QQ Plot Relreturn}

ggplot(t, aes(sample=port1\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port2\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port3\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port4\_Relreturn)) + stat\_qq()

#a = 4500

#c = 7068

#print("Absolute Return in 10 years")

#ggplot(t[a:c,], aes(sample=port1\_absreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port2\_absreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port3\_absreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port4\_absreturn)) + stat\_qq()

```

#QQ plot of Percentage Return

```{r QQ Plot Percentage Return}

ggplot(t, aes(sample=port1\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port2\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port3\_Perreturn)) + stat\_qq()

ggplot(t, aes(sample=port4\_Perreturn)) + stat\_qq()

#a = 4500

#c = 7068

#print("Percentage Return in 10 years")

#ggplot(t[a:c,], aes(sample=port1\_Perreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port2\_Perreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port3\_Perreturn)) + stat\_qq()

#ggplot(t[a:c,], aes(sample=port4\_Perreturn)) + stat\_qq()

```

#Statistics of Relreturn

```{r Statistics of Relreturn}

cat(red("Relreturn Statistics in all time\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Relreturn)

describe(t$port1\_Relreturn)

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Relreturn)

describe(t$port2\_Relreturn)

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Relreturn)

describe(t$port3\_Relreturn)

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Relreturn)

describe(t$port4\_Relreturn)

# a = 4500

# c = 7068

#

# cat(red("absreturn Statistics in 10 years\n"))

# cat(blue("Portfolio 1\n"))

# stat.desc(t$port1\_absreturn[a:c])

# describe(t$port1\_absreturn[a:c])

#

# cat(blue("Portfolio 2\n"))

# stat.desc(t$port2\_absreturn[a:c])

# describe(t$port2\_absreturn[a:c])

#

# cat(blue("Portfolio 3\n"))

# stat.desc(t$port3\_absreturn[a:c])

# describe(t$port3\_absreturn[a:c])

#

# cat(blue("Portfolio 4\n"))

# stat.desc(t$port4\_absreturn[a:c])

# describe(t$port4\_absreturn[a:c])

```

#Statistics of Percentage Return

```{r Statistics of Percentage Return}

cat(red("Percentage Statistics in all time\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Perreturn)

describe(t$port1\_Perreturn)

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Perreturn)

describe(t$port2\_Perreturn)

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Perreturn)

describe(t$port3\_Perreturn)

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Perreturn)

describe(t$port4\_Perreturn)

a = 4500

c = 7068

cat(red("Percentage Statistics in 10 years\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Perreturn[a:c])

describe(t$port1\_Perreturn[a:c])

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Perreturn[a:c])

describe(t$port2\_Perreturn[a:c])

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Perreturn[a:c])

describe(t$port3\_Perreturn[a:c])

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Perreturn[a:c])

describe(t$port4\_Perreturn[a:c])

```

#Statistics Comparison

```{r Statistics table}

base <- t[,c(10,11,12,13)]

colnames(base) <- c("port1\_Relreturn","port2\_Relreturn","port3\_Relreturn","port4\_Relreturn")

#base <- t[,c(9,11,13,15)]

#colnames(base) <- c("port1\_absreturn","port2\_absreturn","port3\_absreturn","port4\_absreturn")

stat <- sapply(base,sum) %>%

rbind(sapply(base,mean)) %>%

rbind(sapply(base,mean)\*252) %>%

rbind(sapply(base,var)) %>%

rbind(sapply(base,skew)) %>%

rbind(sapply(base,kurtosis))

rownames(stat) <- c("sum","mean","annual mean","var","skew","kurtosis")

print("Statistics in all time")

print(stat)

# a = 4500

# c = 7068

#

# stat2 <- sapply(base[a:c,],sum) %>%

# rbind(sapply(base[a:c,],mean)) %>%

# rbind(sapply(base[a:c,],mean)\*252) %>%

# rbind(sapply(base[a:c,],var)) %>%

# rbind(sapply(base[a:c,],skew)) %>%

# rbind(sapply(base[a:c,],kurtosis))

#

# rownames(stat2) <- c("sum","mean","annual mean","var","skew","kurtosis")

#

# print("Statistics in 10 years")

# print(stat2)

```

#Cumulative RelReturn

See how the cost of options and straddles ruined your return

```{r Cumulative Relreturn}

t$port1\_cum\_Relreturn <- cumsum(t$port1\_Relreturn)

t$port2\_cum\_Relreturn <- cumsum(t$port2\_Relreturn)

t$port3\_cum\_Relreturn <- cumsum(t$port3\_Relreturn)

t$port4\_cum\_Relreturn <- cumsum(t$port4\_Relreturn)

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_Relreturn,color="Port1")) +

geom\_line(aes(y=port2\_cum\_Relreturn,color="Port2")) +

geom\_line(aes(y=port3\_cum\_Relreturn,color="Port3")) +

geom\_line(aes(y=port4\_cum\_Relreturn,color="Port4")) +

# ylim(-10000000,50000000) +

theme\_bw() +

ggtitle("Cumulative Return comparing in all time")

# a = 4500

# c = 7068

# t$port1\_cum\_abs\_return2[a:c]<-cumsum(t$port1\_Relreturn[a:c])

# t$port2\_cum\_abs\_return2[a:c]<-cumsum(t$port2\_Relreturn[a:c])

# t$port3\_cum\_abs\_return2[a:c]<-cumsum(t$port3\_Relreturn[a:c])

# t$port4\_cum\_abs\_return2[a:c]<-cumsum(t$port4\_Relreturn[a:c])

#

# ggplot(t[a:c,],aes(x=Date)) +

# geom\_line(aes(y=port1\_cum\_abs\_return2,color="Port1")) +

# geom\_line(aes(y=port2\_cum\_abs\_return2,color="Port2")) +

# geom\_line(aes(y=port3\_cum\_abs\_return2,color="Port3")) +

# geom\_line(aes(y=port4\_cum\_abs\_return2,color="Port4")) +

# theme\_bw() +

# ggtitle("abs Return comparing in 10 years")

```

#Culmulative percentage Return(see initiate invest as $1)

See how the cost of options and straddles ruined your return

```{r Culmulative per returns}

t$port1\_cum\_per\_return<-cumsum(t$port1\_Perreturn)

t$port2\_cum\_per\_return<-cumsum(t$port2\_Perreturn)

t$port3\_cum\_per\_return<-cumsum(t$port3\_Perreturn)

t$port4\_cum\_per\_return<-cumsum(t$port4\_Perreturn)

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_per\_return,color="Port1")) +

geom\_line(aes(y=port2\_cum\_per\_return,color="Port2")) +

geom\_line(aes(y=port3\_cum\_per\_return,color="Port3")) +

geom\_line(aes(y=port4\_cum\_per\_return,color="Port4")) +

theme\_bw() +

ggtitle("per Return comparing in all time")

a = 4500

c = 7068

t$port1\_cum\_per\_return2[a:c]<-cumsum(t$port1\_Perreturn[a:c])

t$port2\_cum\_per\_return2[a:c]<-cumsum(t$port2\_Perreturn[a:c])

t$port3\_cum\_per\_return2[a:c]<-cumsum(t$port3\_Perreturn[a:c])

t$port4\_cum\_per\_return2[a:c]<-cumsum(t$port4\_Perreturn[a:c])

ggplot(t[a:c,],aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_per\_return2,color="Port1")) +

geom\_line(aes(y=port2\_cum\_per\_return2,color="Port2")) +

geom\_line(aes(y=port3\_cum\_per\_return2,color="Port3")) +

geom\_line(aes(y=port4\_cum\_per\_return2,color="Port4")) +

theme\_bw() +

ggtitle("per Return comparing in 10 years")

```

#Realized Volatility on Relreturn

Realized Vol is limited, 10 years has upper boundary

```{r Realized Volatility}

t$port1\_real\_vol <- 252\*cumsum(t$port1\_Relreturn^2)/seq(length(t$port1\_Relreturn)) %>% sqrt()

t$port2\_real\_vol <- 252\*cumsum(t$port2\_Relreturn^2)/seq(length(t$port2\_Relreturn)) %>% sqrt()

t$port3\_real\_vol <- 252\*cumsum(t$port3\_Relreturn^2)/seq(length(t$port3\_Relreturn)) %>% sqrt()

t$port4\_real\_vol <- 252\*cumsum(t$port4\_Relreturn^2)/seq(length(t$port4\_Relreturn)) %>% sqrt()

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_real\_vol,color="Port1")) +

geom\_line(aes(y=port2\_real\_vol,color="Port2")) +

geom\_line(aes(y=port3\_real\_vol,color="Port3")) +

geom\_line(aes(y=port4\_real\_vol,color="Port4")) +

# ylim(-1000,10000) +

theme\_bw() +

ggtitle("Realized Volatility in all time")

# a = 4500

# c = 7068

#

# t$port1\_real\_vol2[a:c] <- 252\*cumsum(t$port1\_absreturn[a:c]^2)/seq(length(t$port1\_absreturn[a:c])) %>% sqrt()

# t$port2\_real\_vol2[a:c] <- 252\*cumsum(t$port2\_absreturn[a:c]^2)/seq(length(t$port2\_absreturn[a:c])) %>% sqrt()

# t$port3\_real\_vol2[a:c] <- 252\*cumsum(t$port3\_absreturn[a:c]^2)/seq(length(t$port3\_absreturn[a:c])) %>% sqrt()

# t$port4\_real\_vol2[a:c] <- 252\*cumsum(t$port4\_absreturn[a:c]^2)/seq(length(t$port4\_absreturn[a:c])) %>% sqrt()

#

# ggplot(t[a:c,],aes(x=Date)) +

# geom\_line(aes(y=port1\_real\_vol2,color="Port1")) +

# geom\_line(aes(y=port2\_real\_vol2,color="Port2")) +

# geom\_line(aes(y=port3\_real\_vol2,color="Port3")) +

# geom\_line(aes(y=port4\_real\_vol2,color="Port4")) +

# theme\_bw() +

# ggtitle("Realized Volatility in 10 years")

```

#Rolling Analysis annually(252 days)

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis annually}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

port4 <- xts(x=t$port4\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,252,gap=252,by=1,FUN="mean")

t$port2\_roll\_mean <- apply.rolling(port2,252,gap=252,by=1,FUN="mean")

t$port3\_roll\_mean <- apply.rolling(port3,252,gap=252,by=1,FUN="mean")

t$port4\_roll\_mean <- apply.rolling(port4,252,gap=252,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,252,gap=252,by=1,FUN="sd")

t$port2\_roll\_vol <- apply.rolling(port2,252,gap=252,by=1,FUN="sd")

t$port3\_roll\_vol <- apply.rolling(port3,252,gap=252,by=1,FUN="sd")

t$port4\_roll\_vol <- apply.rolling(port4,252,gap=252,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,252)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,252)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,252)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,252)

```

#Rolling Analysis 60 days

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis 60 days}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,60,gap=60,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,60,gap=60,by=1,FUN="sd")

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

t$port2\_roll\_mean <- apply.rolling(port2,60,gap=60,by=1,FUN="mean")

t$port2\_roll\_vol <- apply.rolling(port2,60,gap=60,by=1,FUN="sd")

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

t$port3\_roll\_mean <- apply.rolling(port3,60,gap=60,by=1,FUN="mean")

t$port3\_roll\_vol <- apply.rolling(port3,60,gap=60,by=1,FUN="sd")

port4 <- xts(x=t$port4\_Relreturn, order.by=t$Date)

t$port4\_roll\_mean <- apply.rolling(port4,60,gap=60,by=1,FUN="mean")

t$port4\_roll\_vol <- apply.rolling(port4,60,gap=60,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,60)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,60)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,60)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,60)

```

**Distribution**

---

title: "Distribution"

author: "Panther"

date: "February 10, 2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);

# here data read is being processed

#t <- read\_csv("../data/data\_div.csv")

#t <- read\_csv("../Data/data.csv")

t <- read\_csv("F:/Dropbox/[IAQF]/Data/data\_final.csv")

t <- t[-1,] #dump first day NA

# t$Date <- strptime(as.character(t$Date), "%m/%d/%Y")

# format(t$Date, "%Y-%m-%d")

# t$Date <- as.Date(t$Date)

```

#Plot SPC and moving average

```{r Plot index and moving average}

t$POSIXct\_Date <- as.POSIXct(t$Date)

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for all data")

a = 4500

c = 7069

ggplot(t[a:c,],aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

ggtitle("Closing index and moving averge for 10 years")

```

#Momentum effect of portfolio returns

Obvious autocorrelation, large ARMA parameters

```{r Momentum effect}

acf(t$port1\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port1\_Relreturn,lag.max = 50)

acf(t$port2\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port2\_Relreturn,lag.max = 50)

acf(t$port3\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port3\_Relreturn,lag.max = 50)

acf(t$port4\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port4\_Relreturn,lag.max = 50)

```

# Scatter of RelReturn

```{r Scatter Plot RelReturn}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port4\_Relreturn),size =1, shape= 1)

# ylim(-250000,350000)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Relreturn,color="Port1"),alpha=0.2) +

geom\_point(aes(y=port2\_Relreturn,color="Port2"),alpha=0.2) +

geom\_point(aes(y=port3\_Relreturn,color="Port3"),alpha=0.2) +

geom\_point(aes(y=port4\_Relreturn,color="Port4"),alpha=0.2) +

ylim(-100000,200000) +

theme\_bw() +

ggtitle("Relreturn comparing for four portfolios")

```

#Histogram of RelReturn

Portfolio 2 is strongly non-central and skewed

Portfolio 4 is pretty discrete

```{r Histogram Plot RelReturn}

ggplot(t,aes(x = port1\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port1\_Relreturn),sd = sd(t$port1\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port2\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port2\_Relreturn),sd = sd(t$port2\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port3\_Relreturn)) +

geom\_histogram(aes(y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port3\_Relreturn),sd = sd(t$port3\_Relreturn))) +

theme\_bw()

ggplot(t,aes(x = port4\_Relreturn)) +

geom\_histogram(aes(binwidth = 1,y=..density..,fill=..count..)) +

stat\_function(fun = dnorm, color="red", args = list(mean = mean(t$port4\_Relreturn),sd = sd(t$port4\_Relreturn))) +

# xlim(-10000,10000) +

theme\_bw()

```

#QQ plot of Relreturn

```{r QQ Plot Relreturn}

ggplot(t, aes(sample=port1\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port2\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port3\_Relreturn)) + stat\_qq()

ggplot(t, aes(sample=port4\_Relreturn)) + stat\_qq()

```

#Statistics of Relreturn

```{r Statistics of Relreturn}

cat(red("Relreturn Statistics in all time\n"))

cat(blue("Portfolio 1\n"))

stat.desc(t$port1\_Relreturn)

describe(t$port1\_Relreturn)

cat(blue("Portfolio 2\n"))

stat.desc(t$port2\_Relreturn)

describe(t$port2\_Relreturn)

cat(blue("Portfolio 3\n"))

stat.desc(t$port3\_Relreturn)

describe(t$port3\_Relreturn)

cat(blue("Portfolio 4\n"))

stat.desc(t$port4\_Relreturn)

describe(t$port4\_Relreturn)

```

#Statistics Comparison

```{r Statistics table}

base <- t[,c(11,13,15,17)]

colnames(base) <- c("port1\_Relreturn","port2\_Relreturn","port3\_Relreturn","port4\_Relreturn")

#base <- t[,c(9,11,13,15)]

#colnames(base) <- c("port1\_absreturn","port2\_absreturn","port3\_absreturn","port4\_absreturn")

stat <- sapply(base,sum) %>%

rbind(sapply(base,mean)) %>%

rbind(sapply(base,mean)\*252) %>%

rbind(sapply(base,var)) %>%

rbind(sapply(base,skew)) %>%

rbind(sapply(base,kurtosis))

rownames(stat) <- c("sum","mean","annual mean","var","skew","kurtosis")

print("Statistics in all time")

print(stat)

```

#Cumulative RelReturn

See how the cost of options and straddles ruined your return

```{r Cumulative Relreturn}

t$port1\_cum\_Relreturn <- cumsum(t$port1\_Relreturn)

t$port2\_cum\_Relreturn <- cumsum(t$port2\_Relreturn)

t$port3\_cum\_Relreturn <- cumsum(t$port3\_Relreturn)

t$port4\_cum\_Relreturn <- cumsum(t$port4\_Relreturn)

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_cum\_Relreturn,color="Port1")) +

geom\_line(aes(y=port2\_cum\_Relreturn,color="Port2")) +

geom\_line(aes(y=port3\_cum\_Relreturn,color="Port3")) +

geom\_line(aes(y=port4\_cum\_Relreturn,color="Port4")) +

# ylim(-10000000,50000000) +

theme\_bw() +

ggtitle("Cumulative Return comparing in all time")

```

#Realized Volatility on Relreturn

Realized Vol is limited, 10 years has upper boundary

```{r Realized Volatility}

t$port1\_real\_vol <- 252\*cumsum(t$port1\_Relreturn^2)/seq(length(t$port1\_Relreturn)) %>% sqrt()

t$port2\_real\_vol <- 252\*cumsum(t$port2\_Relreturn^2)/seq(length(t$port2\_Relreturn)) %>% sqrt()

t$port3\_real\_vol <- 252\*cumsum(t$port3\_Relreturn^2)/seq(length(t$port3\_Relreturn)) %>% sqrt()

t$port4\_real\_vol <- 252\*cumsum(t$port4\_Relreturn^2)/seq(length(t$port4\_Relreturn)) %>% sqrt()

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_real\_vol,color="Port1")) +

geom\_line(aes(y=port2\_real\_vol,color="Port2")) +

geom\_line(aes(y=port3\_real\_vol,color="Port3")) +

geom\_line(aes(y=port4\_real\_vol,color="Port4")) +

# ylim(-1000,10000) +

theme\_bw() +

ggtitle("Realized Volatility in all time")

```

#Rolling Analysis annually(252 days)

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis annually}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

port4 <- xts(x=t$port4\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,252,gap=252,by=1,FUN="mean")

t$port2\_roll\_mean <- apply.rolling(port2,252,gap=252,by=1,FUN="mean")

t$port3\_roll\_mean <- apply.rolling(port3,252,gap=252,by=1,FUN="mean")

t$port4\_roll\_mean <- apply.rolling(port4,252,gap=252,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,252,gap=252,by=1,FUN="sd")

t$port2\_roll\_vol <- apply.rolling(port2,252,gap=252,by=1,FUN="sd")

t$port3\_roll\_vol <- apply.rolling(port3,252,gap=252,by=1,FUN="sd")

t$port4\_roll\_vol <- apply.rolling(port4,252,gap=252,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,252)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,252)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,252)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,252)

```

#Rolling Analysis 60 days

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis 60 days}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,60,gap=60,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,60,gap=60,by=1,FUN="sd")

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

t$port2\_roll\_mean <- apply.rolling(port2,60,gap=60,by=1,FUN="mean")

t$port2\_roll\_vol <- apply.rolling(port2,60,gap=60,by=1,FUN="sd")

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

t$port3\_roll\_mean <- apply.rolling(port3,60,gap=60,by=1,FUN="mean")

t$port3\_roll\_vol <- apply.rolling(port3,60,gap=60,by=1,FUN="sd")

port4 <- xts(x=t$port4\_Relreturn, order.by=t$Date)

t$port4\_roll\_mean <- apply.rolling(port4,60,gap=60,by=1,FUN="mean")

t$port4\_roll\_vol <- apply.rolling(port4,60,gap=60,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port4\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,60)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,60)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,60)

chart.RollingCorrelation(t$port4\_roll\_mean, t$port4\_roll\_vol,60)

```

**Portfolio Structuring 2**

---

title: "Portfolio Structuring"

author: "Panther"

date: "February 9, 2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);

# here data read is being processed

setwd("F:/Dropbox/[IAQF\_Max]/R")

#t <- read\_csv("../data/data\_div.csv")

#t <- read\_csv("../Data/data.csv")

#t <- read\_csv("F:/Dropbox/[IAQF]/Data/data.csv")

# t$Date <- strptime(as.character(t$Date), "%m/%d/%Y")

# format(t$Date, "%Y-%m-%d")

# t$Date <- as.Date(t$Date)

```

#Process Data

```{r Process Data}

GSPC <- read\_csv("../data/^GSPC.csv")

Treasury <- read\_csv("../data/DTB3.csv")

VIX <- read\_csv("../data/^VIX.csv")

DIV <- read\_csv("../data/Dividends.csv")

DIV\_M <- read\_csv("../data/SP500\_DIV\_YIELD\_MONTH.csv")

# DIV$`SPX Index` <- strptime(as.character(DIV$`SPX Index`), "%m/%d/%Y")

# format(DIV$`SPX Index`, "%Y-%m-%d")

# DIV$`SPX Index` <- as.Date(DIV$`SPX Index`)

# t <- t[-10]

t<-cbind(GSPC,Treasury$DTB3[match(GSPC$Date,Treasury$DATE)])

t<-cbind(t,VIX$Close[match(t$Date,VIX$Date)])

t<-cbind(t,DIV$X2[match(t$Date,DIV$`SPX Index`)])

t<-t[,c(1,5,8,9,10)]

t<- t[-1:(-10054+120),] # extra 120 for training signal, will remove later

colnames(t)[3] <- "Free\_rate"

colnames(t)[4] <- "VIX"

colnames(t)[5] <- "Div\_rate\_daily"

t$Free\_rate = as.numeric(as.character(t$Free\_rate))

t$Div\_rate\_daily = as.numeric(as.character(t$Div\_rate\_daily))

```

#Missing Data

```{r Missing Data}

sum(is.na(t))

sum(is.na(t$Close))

sum(is.na(t$Free\_rate))

sum(is.na(t$VIX))

sum(is.na(t$Div\_rate\_daily))

sum(is.na(DIV\_M))

which(is.na(t$Free\_rate))

#count(which(is.na(t$Free\_rate)))

which(is.na(t$VIX)) #don't worry, this missing 120 is not needed

which(is.na(t$Div\_rate\_daily))

#missing data of free rate

temp <- which(is.na(t$Free\_rate))

for (i in temp)

t$Free\_rate[i] <- mean(t$Free\_rate[(i-5):(i+5)], na.rm=TRUE)

t$Div\_rate\_daily[2] <- mean(t$Div\_rate\_daily[1:32], na.rm=TRUE)

t$Div\_rate\_daily[7] <- mean(t$Div\_rate\_daily[1:32], na.rm=TRUE)

temp <- which(is.na(t$Div\_rate\_daily))

for (i in temp)

t$Div\_rate\_daily[i] <- mean(t$Div\_rate\_daily[(i-15):(i+15)], na.rm=TRUE)

sum(is.na(t$Free\_rate))

sum(is.na(t$Div\_rate\_daily))

```

#Smooth dividend daily rate as monthly rate

```{r smooth dividend}

t$Div\_rate\_monthly <- t$Div\_rate\_daily

# t$POSIXct\_Date <- as.POSIXct(t$Date)

# DIV\_M$POSIXct\_Date <- as.POSIXct(DIV\_M$Date)

DIV\_M <- DIV\_M[ nrow(DIV\_M):1, ]

DIV\_M<- DIV\_M[-1:-1416,]

temp <- c("2018-01-31",1.87)

DIV\_M <- rbind(DIV\_M, temp)

DIV\_M$Value = as.numeric(DIV\_M$Value)

#

# for (i in seq\_along(t$Div\_rate\_daily)){

# if (i < 16)

# t$Div\_rate\_monthly[i] = mean(t$Div\_rate\_daily[i:(i+30)], na.rm=TRUE)

# else

# t$Div\_rate\_monthly[i] = mean(t$Div\_rate\_daily[(i-15):(i+15)], na.rm=TRUE)}

#

# ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=Div\_rate\_daily))

# ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=Div\_rate\_monthly))

#

# stat.desc(t$Div\_rate\_daily)

# describe(t$Div\_rate\_daily)

# stat.desc(t$Div\_rate\_monthly)

# describe(t$Div\_rate\_monthly)

for (i in seq\_along(t$Div\_rate\_daily)){

for (j in seq\_along(DIV\_M$Value)){

if (t$Date[i] <= DIV\_M$Date[j+1] && t$Date[i] > DIV\_M$Date[j]){

t$Div\_rate\_monthly[i] = DIV\_M$Value[j+1]/100

break

}

}

}

```

#Moving Average and momentum signal

```{r momentum signal}

t$SMA60 <- SMA(t$Close,60)

t$SMA120 <- SMA(t$Close,120)

t<- t[-1:-120,] # here remove training 120

for (i in seq\_along(t$SMA120)){

if (t$SMA60[i]>t$SMA120[i])

t$signal[i] = 1 # signal = 1 is buy singal

else if (t$SMA60[i]<t$SMA120[i])

t$signal[i] = 0

else

t$signal[i] = t$singal[i-1]

}

```

## Create portfolios

Absolute Return = shift between one lag of close prices

Percentage Return = Absolute Return / last day close price

Relative Return = Percentage return \* initiate price($100,000)

We use Relative Return as analysis object

#portfolio 1

Buy/short 1 index on today's closing and close position next day

Assume when you short today, you buy an index for paying back next day

Recieve dividend while buying

```{r portfolio 1}

t$port1\_absreturn[1] = NA

#t$port1\_Perreturn[1] = NA

t$port1\_Relreturn[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

t$port1\_absreturn[i+1] = (t$Close[i+1]-t$Close[i]+t$Close[i]\*t$Div\_rate\_monthly[i]/365\*as.numeric(difftime(t$Date[i+1],t$Date[i])))

# t$port1\_Perreturn[i+1] = (t$Close[i+1]-t$Close[i]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i+1]-t$Close[i]+t$Close[i]\*t$Div\_rate\_monthly[i]/365\*as.numeric(difftime(t$Date[i+1],t$Date[i]))) / t$Close[i] \* 100000

}

else{

t$port1\_absreturn[i+1] = t$Close[i]-t$Close[i+1]

# t$port1\_Perreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i] \* 100000

}

}

t$port1\_Perreturn <- t$port1\_Relreturn/100000

```

#portfolio 2

Buy at the money call/put option(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 2}

t$port2\_absreturn[1] = NA

#t$port2\_Perreturn[1] = NA

t$port2\_Relreturn[1] = NA

#if (t$signal[1] == 1) {

# cc <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

# } else {

# cc <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

# for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

#

# if (t$signal[i]== 1){

# aa <- GBSOption("c",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

# bb <- GBSOption("c",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

# } else{

# aa <- GBSOption("p",t$Close[i+1],t$Close[i],(90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365,(t$Free\_rate[i+1]/100),0,(t$VIX[i+1]/100))

# bb <- GBSOption("p",t$Close[i],t$Close[i],0.246575342,(t$Free\_rate[i]/100),0,(t$VIX[i]/100))

# }

# t$port2\_absreturn[i+1] = aa@price-bb@price

# t$port2\_Perreturn[i+1] = (aa@price-bb@price) / bb@price

# # t$port2\_Relreturn[i+1] = t$port2\_Perreturn[i+1] \* t$Close[1]

# }

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

aa <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("call", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

} else {

aa <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("put", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

}

t$port2\_absreturn[i+1] = aa$value-bb$value

# t$port2\_Perreturn[i+1] = (aa$value-bb$value) / bb$value

t$port2\_Relreturn[i+1] = (aa$value-bb$value) / bb$value \* 100000

}

t$port2\_Perreturn <- t$port2\_Relreturn/100000

```

#portfolio 3

Buy at the money straddle(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 3}

t$port3\_absreturn[1] = NA

#t$port3\_Perreturn[1] = NA

t$port3\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

aa1 <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

aa2 <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (90-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb1 <- EuropeanOption("call", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

bb2 <- EuropeanOption("put", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342, (t$VIX[i]/100))

t$port3\_absreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value)

# t$port3\_Perreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)

t$port3\_Relreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)\*100000

}

t$port3\_Perreturn <- t$port3\_Relreturn/100000

```

#portfolio 4

Buy at the money straddle(90 days maturity but remain life decrease until rebalance at maturity)

- on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 4}

#t$port4\_absreturn[1] = NA

#t$port4\_Perreturn[1] = NA

t$port4\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

#date.lookup <- format(seq(as.Date("2000-01-02"), as.Date("2018-1-19"), by = "1 day"))

#date.except <- subset(b, !(y %in% a$x))

#match("2024-01-19", date.lookup)

count=0 #count days passed

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (as.numeric(difftime(t$Date[i+1],t$Date[i])) == 1){ #workday

count = count+1}

else { # not workday

count = count+as.numeric(difftime(t$Date[i+1],t$Date[i])) }

if (count <= 89){

Remaina = 90 - count # today's remain

Remainb = Remaina + 1} # previous day's remain

else if (count >= 90) { # expired already, adjuste to 90 days passed

Remaina = 0 # today's remain

Remainb = 1 # previous day's remain

count = 0 } # after expired, rebalance

aa1 <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), 0.246575342\*Remaina/90, (t$VIX[i+1]/100),0, 0)

aa2 <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), 0.246575342\*Remaina/90, (t$VIX[i+1]/100),0, 0)

bb1 <- EuropeanOption("call", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342\*Remainb/90, (t$VIX[i]/100),0, 0)

bb2 <- EuropeanOption("put", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 0.246575342\*Remainb/90, (t$VIX[i]/100),0, 0)

# t$port4\_absreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value)

# t$port4\_Perreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)

t$port4\_Relreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)\*100000

}

t$port4\_Perreturn <- t$port4\_Relreturn/100000

```

#Dump first day NA return

```{r}

#t <- t[-1,]

```

**Portfolio 1 and 2 Analysis**

---

title: "Port1&2"

author: "Panther"

date: "February 16, 2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);library(ggrepel);

# here data read is being processed

t <- read\_csv("F:/Dropbox/[IAQF]/Data/data\_final.csv")

t <- t[-1,] #dump first day NA

```

##Signal

#Singal turning point

```{r }

#turning point

t$signal\_turning[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i+1] == t$signal[i]) #different from previous day

t$signal\_turning[i+1] = 0

else

t$signal\_turning[i+1] = 1 #change to and maintain

}

table(t$signal\_turning)

temp = which(t$signal\_turning == 1)

j = 1

a <- data.frame(date=as.POSIXct(character()),close=as.numeric(),signal=as.numeric())

for (i in temp){

a[j,1] <- t$Date[i]

a[j,2] <- t$Close[i]

a[j,3] <- t$signal[i]

j = j+1

}

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

geom\_line(aes(y=SMA60,color="SMA60")) +

geom\_line(aes(y=SMA120,color="SMA120"))+

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")+

scale\_colour\_manual("", breaks = c("Close", "SMA60","SMA120"),

values = c("blue", "red","green")) +

# geom\_point(data=a,aes(x=date,y=close),size =3, shape= 5) +

# geom\_text(data=a, aes(x=date,y=close,label = signal),size = 5,check\_overlap = TRUE, vjust = 0.5) +

# geom\_text\_repel(data=a, aes(x=date,y=close,label = signal)) +

geom\_label\_repel(data=a, aes(x=date,y=close,label = signal),alpha = 0.3) +

geom\_vline(data=a, aes(xintercept =date),linetype="dotted", size=0.2) +

theme\_bw() +

ggtitle("Closing index and moving averge for all data")

```

#Profitability of each turning signal

```{r }

#t <- t[-1,]

mean\_return = mean(t$port1\_Relreturn)

b <- data.frame(sum\_return=as.numeric(),excess\_return=as.numeric(),return\_ratio=as.numeric(),date=as.Date(character()))

count = 1

for (i in temp){

if (count == 1){

b[count,1] <- sum(t$port1\_Relreturn[1:i])

b[count,2] <- b[count,1] - mean\_return \* i

b[count,3] <- b[count,1] / (mean\_return \* i)

b[count,4] <- t$Date[1]

temp1 = i

count = count +1}

else {

b[count,1] <- sum(t$port1\_Relreturn[temp1:i])

b[count,2] <- b[count,1] - mean\_return \* (i-temp1)

b[count,3] <- b[count,1] / (mean\_return \* (i-temp1))

b[count,4] <- t$Date[temp1]

temp1 = i

count = count +1}

}

b[count,1] <- sum(t$port1\_Relreturn[temp1:length(t$port1\_Relreturn)])

b[count,2] <- b[count,1] - mean\_return \* (length(t$port1\_Relreturn)-temp1)

b[count,3] <- b[count,1] / (mean\_return \* (length(t$port1\_Relreturn)-temp1))

b[count,4] <- t$Date[temp1]

which(b$sum\_return >0)

b$row <- as.numeric(row.names(b))

ggplot(b,aes(row,sum\_return)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("sum return")

ggplot(b,aes(row,excess\_return)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("excess return")

ggplot(b,aes(row,return\_ratio)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("return ratio")

aa = 0

bb = 0

for (i in b$row) {

if (i==length(b$row)){

if (b$sum\_return[i]>0)

aa = aa + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))

else

bb = bb + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))}

else {

if (b$sum\_return[i]>0)

aa = aa + as.numeric(difftime(b$date[i+1],b$date[i]))

else

bb = bb + as.numeric(difftime(b$date[i+1],b$date[i]))}

}

aa

bb

aaa = 0

bbb = 0

for (i in b$row) {

if (i==length(b$row)){

if (b$excess\_return[i]>0)

aaa = aaa + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))

else

bbb = bbb + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))}

else {

if (b$excess\_return[i]>0)

aaa = aaa + as.numeric(difftime(b$date[i+1],b$date[i]))

else

bbb = bbb + as.numeric(difftime(b$date[i+1],b$date[i]))}

}

aaa

bbb

```

#Quantify signals' profitability on 60 days rolling window

```{r }

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,60,gap=60,by=1,FUN="mean")

rolling\_mean = mean(t$port1\_roll\_mean,na.rm=TRUE)

c <- data.frame(return\_60days=as.numeric(),excess\_return=as.numeric(),return\_ratio=as.numeric(),date=as.Date(character()))

for (i in seq\_along(t$Close[1:(length(t$Close)-60)])){

c[i,1] <- sum(t$port1\_Relreturn[i:i+59])

c[i,2] <- c[i,1] - rolling\_mean\*60

c[i,3] <- c[i,1] / (rolling\_mean\*60)

c[i,4] <- t$Date[i]

}

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_60days),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=excess\_return),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_ratio),size =1, shape= 1)

cc = 0

dd = 0

cc\_sum = 0

dd\_sum = 0

for (i in seq\_along(c$return\_60days)) {

if (c$return\_60days[i]>0){

cc = cc + 1

cc\_sum = cc\_sum + c$return\_60days[i]}

else {

dd = dd + 1

dd\_sum = dd\_sum + c$return\_60days[i]}

}

cc

dd

cc\_sum

dd\_sum

cc/cc\_sum

dd/dd\_sum

```

#Quantify signals' profitability on 252 days rolling window

```{r }

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,252,gap=252,by=1,FUN="mean")

rolling\_mean = mean(t$port1\_roll\_mean,na.rm=TRUE)

c <- data.frame(return\_60days=as.numeric(),excess\_return=as.numeric(),return\_ratio=as.numeric(),date=as.Date(character()))

for (i in seq\_along(t$Close[1:(length(t$Close)-252)])){

c[i,1] <- sum(t$port1\_Relreturn[i:i+251])

c[i,2] <- c[i,1] - rolling\_mean\*252

c[i,3] <- c[i,1] / (rolling\_mean\*252)

c[i,4] <- t$Date[i]

}

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_60days),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=excess\_return),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_ratio),size =1, shape= 1)

cc = 0

dd = 0

cc\_sum = 0

dd\_sum = 0

for (i in seq\_along(c$return\_60days)) {

if (c$return\_60days[i]>0){

cc = cc + 1

cc\_sum = cc\_sum + c$return\_60days[i]}

else {

dd = dd + 1

dd\_sum = dd\_sum + c$return\_60days[i]}

}

cc

dd

cc\_sum

dd\_sum

cc/cc\_sum

dd/dd\_sum

```

#Profitability of each turning signal ( portfolio 2)

```{r }

#t <- t[-1,]

mean\_return = mean(t$port2\_Relreturn)

b <- data.frame(sum\_return=as.numeric(),excess\_return=as.numeric(),return\_ratio=as.numeric(),date=as.Date(character()))

count = 1

for (i in temp){

if (count == 1){

b[count,1] <- sum(t$port2\_Relreturn[1:i])

b[count,2] <- b[count,1] - mean\_return \* i

b[count,3] <- b[count,1] / (mean\_return \* i)

b[count,4] <- t$Date[1]

temp1 = i

count = count +1}

else {

b[count,1] <- sum(t$port2\_Relreturn[temp1:i])

b[count,2] <- b[count,1] - mean\_return \* (i-temp1)

b[count,3] <- b[count,1] / (mean\_return \* (i-temp1))

b[count,4] <- t$Date[temp1]

temp1 = i

count = count +1}

}

b[count,1] <- sum(t$port2\_Relreturn[temp1:length(t$port2\_Relreturn)])

b[count,2] <- b[count,1] - mean\_return \* (length(t$port2\_Relreturn)-temp1)

b[count,3] <- b[count,1] / (mean\_return \* (length(t$port2\_Relreturn)-temp1))

b[count,4] <- t$Date[temp1]

which(b$sum\_return >0)

b$row <- as.numeric(row.names(b))

ggplot(b,aes(row,sum\_return)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("sum return")

ggplot(b,aes(row,excess\_return)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("excess return")

ggplot(b,aes(row,return\_ratio)) +

geom\_bar(stat = "identity",width=1) +

geom\_label\_repel(data=b, aes(x=row,label=as.character(date)),size = 3,alpha = 0.5) +

ggtitle("return ratio")

aa = 0

bb = 0

for (i in b$row) {

if (i==length(b$row)){

if (b$sum\_return[i]>0)

aa = aa + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))

else

bb = bb + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))}

else {

if (b$sum\_return[i]>0)

aa = aa + as.numeric(difftime(b$date[i+1],b$date[i]))

else

bb = bb + as.numeric(difftime(b$date[i+1],b$date[i]))}

}

aa

bb

aaa = 0

bbb = 0

for (i in b$row) {

if (i==length(b$row)){

if (b$excess\_return[i]>0)

aaa = aaa + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))

else

bbb = bbb + as.numeric(difftime(t$Date[length(t$Date)],b$date[i]))}

else {

if (b$excess\_return[i]>0)

aaa = aaa + as.numeric(difftime(b$date[i+1],b$date[i]))

else

bbb = bbb + as.numeric(difftime(b$date[i+1],b$date[i]))}

}

aaa

bbb

```

#Quantify signals' profitability on 60 days rolling window (Portfolio 2)

```{r }

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

t$port2\_roll\_mean <- apply.rolling(port2,60,gap=60,by=1,FUN="mean")

rolling\_mean = mean(t$port2\_roll\_mean,na.rm=TRUE)

c <- data.frame(return\_60days=as.numeric(),excess\_return=as.numeric(),return\_ratio=as.numeric(),date=as.Date(character()))

for (i in seq\_along(t$Close[1:(length(t$Close)-60)])){

c[i,1] <- sum(t$port2\_Relreturn[i:i+59])

c[i,2] <- c[i,1] - rolling\_mean\*60

c[i,3] <- c[i,1] / (rolling\_mean\*60)

c[i,4] <- t$Date[i]

}

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_60days),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=excess\_return),size =1, shape= 1)

ggplot(c,aes(x=date)) +

geom\_point(aes(y=return\_ratio),size =1, shape= 1)

cc = 0

dd = 0

cc\_sum = 0

dd\_sum = 0

for (i in seq\_along(c$return\_60days)) {

if (c$return\_60days[i]>0){

cc = cc + 1

cc\_sum = cc\_sum + c$return\_60days[i]}

else {

dd = dd + 1

dd\_sum = dd\_sum + c$return\_60days[i]}

}

cc

dd

cc\_sum

dd\_sum

cc/cc\_sum

dd/dd\_sum

```

##Portfolio2 - Option

#Price level

```{r }

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Close,color="Close")) +

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")

S <- xts(x=t$Close, order.by=t$Date)

q <- xts(x=t$Div\_rate\_monthly, order.by=t$Date)

r <- xts(x=t$Free\_rate, order.by=t$Date)

vix <- xts(x=t$VIX, order.by=t$Date)

T = 90

S1 <- apply.rolling(S,252,gap=0,by=252,FUN="mean")

q1 <- apply.rolling(q,252,gap=0,by=252,FUN="mean")

r1 <- apply.rolling(r,252,gap=0,by=252,FUN="mean")

vix1 <- apply.rolling(vix,252,gap=0,by=252,FUN="mean")

S2 <- S1[which(!is.na(S1))]

q2 <- q1[which(!is.na(q1))]

r2 <- r1[which(!is.na(r1))]/100

vix2 <- vix1[which(!is.na(vix1))]/100

T2 <- T/365

#C <- EuropeanOption("call", S2, S2, q2, r2, T2, vix2)

#P <- EuropeanOption("put", S2, S2, q2, r2, T2, vix2)

C <- list()

P <- list()

for (i in seq\_along(S2)){

C[[i]] <- EuropeanOption("call", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

P[[i]] <- EuropeanOption("put", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

}

#date <- data.frame(date=index(S2))

table1 <- data.frame(date=as.Date(character()),index=as.numeric(),callPrice=as.numeric(),callDelta=as.numeric(),putPrice=as.numeric(),putDelta=as.numeric())

for (i in seq\_along(S2)){

table1[i,1] <- as.Date(index(S2[i]))

table1[i,2] <- S2[[i]]

table1[i,3] <- C[[i]]$value

table1[i,4] <- C[[i]]$delta

table1[i,5] <- P[[i]]$value

table1[i,6] <- P[[i]]$delta

}

table1[,7] <- table1[,2]/table1[,3]\*table1[,4]

table1[,8] <- table1[,2]/table1[,5]\*table1[,6]

colnames(table1)[7] <- "callPortDelta"

colnames(table1)[8] <- "putPortDelta"

print(table1)

t$port2\_cum\_Relreturn <- cumsum(t$port2\_Relreturn)

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=t$port2\_cum\_Relreturn,color="Cum\_port2\_rel")) +

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")

ggplot(table1,aes(x=date)) +

# geom\_line(aes(y=index,color="index")) +

geom\_line(aes(y=callPortDelta,color="call port delta")) +

geom\_line(aes(y=putPortDelta,color="put port delta")) +

labs(color="Legend")

#C[[i]]$

# value delta gamma vega theta rho divRho

```

#VIX

```{r }

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=VIX,color="VIX")) +

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")

ggplot(table1,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=VIX,color="VIX")) +

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")

table2 <- data.frame(date=as.Date(character()),VIX=as.numeric(),callPrice=as.numeric(),callVega=as.numeric(),putPrice=as.numeric(),putVega=as.numeric())

for (i in seq\_along(S2)){

table2[i,1] <- as.Date(index(S2[i]))

table2[i,2] <- vix2[[i]]

table2[i,3] <- C[[i]]$value

table2[i,4] <- C[[i]]$vega

table2[i,5] <- P[[i]]$value

table2[i,6] <- P[[i]]$vega

}

ggplot(table2,aes(x=date)) +

geom\_line(aes(y=callVega,color="Vega"))

```

#r & Div

```{r }

ggplot(t,aes(x=POSIXct\_Date)) +

geom\_line(aes(y=Free\_rate/100,color="r")) +

geom\_line(aes(y=Div\_rate\_monthly,color="q")) +

scale\_x\_datetime(breaks=date\_breaks("1 year"),labels = date\_format("%y")) +

labs(color="Legend")

table3 <- data.frame(date=as.Date(character()),r=as.numeric(),callRho=as.numeric(),putRho=as.numeric(),q=as.numeric(),callDivRho=as.numeric(),putDivRho=as.numeric())

for (i in seq\_along(S2)){

table3[i,1] <- as.Date(index(S2[i]))

table3[i,2] <- r2[[i]]

table3[i,3] <- C[[i]]$rho

table3[i,4] <- P[[i]]$rho

table3[i,5] <- q2[[i]]

table3[i,6] <- P[[i]]$divRho

table3[i,7] <- P[[i]]$divRho

}

ggplot(table3,aes(x=date)) +

geom\_line(aes(y=callRho,color="Rho")) +

geom\_line(aes(y=callDivRho,color="DivRho"))

```

#T

```{r }

table4 <- data.frame(date=as.Date(character()),callTheta=as.numeric(),putTheta=as.numeric())

for (i in seq\_along(S2)){

table4[i,1] <- as.Date(index(S2[i]))

table4[i,2] <- C[[i]]$theta

table4[i,3] <- C[[i]]$theta

}

ggplot(table4,aes(x=date)) +

geom\_line(aes(y=callTheta,color="callTheta")) +

geom\_line(aes(y=putTheta,color="putTheta"))

T = 180

T2 <- T/365

#C <- EuropeanOption("call", S2, S2, q2, r2, T2, vix2)

#P <- EuropeanOption("put", S2, S2, q2, r2, T2, vix2)

C <- list()

P <- list()

for (i in seq\_along(S2)){

C[[i]] <- EuropeanOption("call", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

P[[i]] <- EuropeanOption("put", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

}

for (i in seq\_along(S2)){

table4[i,1] <- as.Date(index(S2[i]))

table4[i,2] <- C[[i]]$theta

table4[i,3] <- C[[i]]$theta

}

ggplot(table4,aes(x=date)) +

geom\_line(aes(y=callTheta,color="callTheta")) +

geom\_line(aes(y=putTheta,color="putTheta"))

T = 45

T2 <- T/365

#C <- EuropeanOption("call", S2, S2, q2, r2, T2, vix2)

#P <- EuropeanOption("put", S2, S2, q2, r2, T2, vix2)

C <- list()

P <- list()

for (i in seq\_along(S2)){

C[[i]] <- EuropeanOption("call", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

P[[i]] <- EuropeanOption("put", S2[i], S2[i], q2[i], r2[i], T2, vix2[i])

}

for (i in seq\_along(S2)){

table4[i,1] <- as.Date(index(S2[i]))

table4[i,2] <- C[[i]]$theta

table4[i,3] <- C[[i]]$theta

}

ggplot(table4,aes(x=date)) +

geom\_line(aes(y=callTheta,color="callTheta")) +

geom\_line(aes(y=putTheta,color="putTheta"))

```

#factor model on all components

```{r }

t$VIX <- t$VIX/100

t$Free\_rate <- t$Free\_rate/100

t$signal <- factor(t$signal)

tt <- cbind(t$port2\_Relreturn,t$Close,t$signal,t$VIX,t$Free\_rate,t$Div\_rate\_monthly)

colnames(tt) <- c("return","close","signal","VIX","Free\_rate","Div")

tt <- data.frame(tt)

Mod1 <- lm(return ~ close+signal+VIX+Free\_rate+Div, data=tt)

summary(Mod1)

lmMod <- lm(return ~ . , data = tt)

selectedMod <- step(lmMod)

summary(selectedMod)

selectedMod <- lm(return ~ close+signal+VIX+Free\_rate+Div, data=tt)

all\_vifs <- car::vif(selectedMod)

print(all\_vifs)

signif\_all <- names(all\_vifs)

# Remove vars with VIF> 4 and re-build model until none of VIFs don't exceed 4.

while(any(all\_vifs > 4)){

var\_with\_max\_vif <- names(which(all\_vifs == max(all\_vifs))) # get the var with max vif

signif\_all <- signif\_all[!(signif\_all) %in% var\_with\_max\_vif] # remove

myForm <- as.formula(paste("return ~ ", paste (signif\_all, collapse=" + "), sep="")) # new formula

selectedMod <- lm(myForm, data=inputData) # re-build model with new formula

all\_vifs <- car::vif(selectedMod)

}

summary(selectedMod)

all\_vars <- names(selectedMod[[1]])[-1] # names of all X variables

# Get the non-significant vars

summ <- summary(selectedMod) # model summary

pvals <- summ[[4]][, 4] # get all p values

not\_significant <- character() # init variables that aren't statsitically significant

not\_significant <- names(which(pvals > 0.1))

not\_significant <- not\_significant[!not\_significant %in% "(Intercept)"] # remove 'intercept'. Optional!

# If there are any non-significant variables,

while(length(not\_significant) > 0){

all\_vars <- all\_vars[!all\_vars %in% not\_significant[1]]

myForm <- as.formula(paste("return ~ ", paste (all\_vars, collapse=" + "), sep="")) # new formula

selectedMod <- lm(myForm, data=tt) # re-build model with new formula

# Get the non-significant vars.

summ <- summary(selectedMod)

pvals <- summ[[4]][, 4]

not\_significant <- character()

not\_significant <- names(which(pvals > 0.1))

not\_significant <- not\_significant[!not\_significant %in% "(Intercept)"]

}

summary(selectedMod)

library(leaps)

regsubsetsObj <- regsubsets(x=tt[,2:6] ,y=tt$return, nbest = 2, really.big = T)

plot(regsubsetsObj, scale = "adjr2") # regsubsets plot based on R-sq

Mod2 <- lm(return ~ close, data=tt)

summary(Mod2)

```

#factor model on all components(monthly)

```{r }

ttt=xts(tt[,2:7], order.by=tt$`t$Date`)

tttt=daily2monthly(ttt, FUN=mean, na.rm=TRUE)

#date <- index(tttt)

tttt <- data.frame(tttt)

Mod1 <- lm(return ~ close+signal+VIX+Free\_rate+Div, data=tttt)

summary(Mod1)

lmMod <- lm(return ~ . , data = tttt)

selectedMod <- step(lmMod)

summary(selectedMod)

selectedMod <- lm(return ~ close+signal+VIX+Free\_rate+Div, data=tt)

all\_vifs <- car::vif(selectedMod)

print(all\_vifs)

signif\_all <- names(all\_vifs)

# Remove vars with VIF> 4 and re-build model until none of VIFs don't exceed 4.

while(any(all\_vifs > 4)){

var\_with\_max\_vif <- names(which(all\_vifs == max(all\_vifs))) # get the var with max vif

signif\_all <- signif\_all[!(signif\_all) %in% var\_with\_max\_vif] # remove

myForm <- as.formula(paste("return ~ ", paste (signif\_all, collapse=" + "), sep="")) # new formula

selectedMod <- lm(myForm, data=inputData) # re-build model with new formula

all\_vifs <- car::vif(selectedMod)

}

summary(selectedMod)

all\_vars <- names(selectedMod[[1]])[-1] # names of all X variables

# Get the non-significant vars

summ <- summary(selectedMod) # model summary

pvals <- summ[[4]][, 4] # get all p values

not\_significant <- character() # init variables that aren't statsitically significant

not\_significant <- names(which(pvals > 0.1))

not\_significant <- not\_significant[!not\_significant %in% "(Intercept)"] # remove 'intercept'. Optional!

# If there are any non-significant variables,

while(length(not\_significant) > 0){

all\_vars <- all\_vars[!all\_vars %in% not\_significant[1]]

myForm <- as.formula(paste("return ~ ", paste (all\_vars, collapse=" + "), sep="")) # new formula

selectedMod <- lm(myForm, data=tt) # re-build model with new formula

# Get the non-significant vars.

summ <- summary(selectedMod)

pvals <- summ[[4]][, 4]

not\_significant <- character()

not\_significant <- names(which(pvals > 0.1))

not\_significant <- not\_significant[!not\_significant %in% "(Intercept)"]

}

summary(selectedMod)

library(leaps)

regsubsetsObj <- regsubsets(x=tttt[,2:6] ,y=tttt$return, nbest = 2, really.big = T)

plot(regsubsetsObj, scale = "adjr2") # regsubsets plot based on R-sq

Mod2 <- lm(return ~ close, data=tt)

summary(Mod2)

Mod2 <- lm(return ~ signal, data=tttt)

summary(Mod2)

```

**Improvement**

---

title: "Improve"

author: "Panther"

date: "February 19, 2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);library(ggrepel);

# here data read is being processed

#t <- read\_csv("F:/Dropbox/[IAQF]/Data/data\_final.csv")

#t <- t[-1,] #dump first day NA

#setwd("F:/Dropbox/[IAQF\_Max]/R")

```

#Process Data

```{r Process Data}

GSPC <- read\_csv("../data/^GSPC.csv")

Treasury <- read\_csv("../data/DTB3.csv")

VIX <- read\_csv("../data/^VIX.csv")

DIV <- read\_csv("../data/Dividends.csv")

DIV\_M <- read\_csv("../data/SP500\_DIV\_YIELD\_MONTH.csv")

# DIV$`SPX Index` <- strptime(as.character(DIV$`SPX Index`), "%m/%d/%Y")

# format(DIV$`SPX Index`, "%Y-%m-%d")

# DIV$`SPX Index` <- as.Date(DIV$`SPX Index`)

# t <- t[-10]

t<-cbind(GSPC,Treasury$DTB3[match(GSPC$Date,Treasury$DATE)])

t<-cbind(t,VIX$Close[match(t$Date,VIX$Date)])

t<-cbind(t,DIV$X2[match(t$Date,DIV$`SPX Index`)])

t<-t[,c(1,5,8,9,10)]

t<- t[-1:(-10054+251),] # extra 120 for training signal, will remove later

colnames(t)[3] <- "Free\_rate"

colnames(t)[4] <- "VIX"

colnames(t)[5] <- "Div\_rate\_daily"

t$Free\_rate = as.numeric(as.character(t$Free\_rate))

t$Div\_rate\_daily = as.numeric(as.character(t$Div\_rate\_daily))

```

#Missing Data

```{r Missing Data}

#missing data of free rate

temp <- which(is.na(t$Free\_rate))

for (i in temp)

t$Free\_rate[i] <- mean(t$Free\_rate[(i-5):(i+5)], na.rm=TRUE)

#t$Div\_rate\_daily[2] <- mean(t$Div\_rate\_daily[1:32], na.rm=TRUE)

#t$Div\_rate\_daily[7] <- mean(t$Div\_rate\_daily[1:32], na.rm=TRUE)

temp <- which(is.na(t$Div\_rate\_daily))

for (i in temp){

if (i<15)

t$Div\_rate\_daily[i] <- mean(t$Div\_rate\_daily[1:30], na.rm=TRUE)

else

t$Div\_rate\_daily[i] <- mean(t$Div\_rate\_daily[(i-15):(i+15)], na.rm=TRUE)}

```

#Smooth dividend daily rate as monthly rate

```{r smooth dividend}

t$Div\_rate\_monthly <- t$Div\_rate\_daily

# t$POSIXct\_Date <- as.POSIXct(t$Date)

# DIV\_M$POSIXct\_Date <- as.POSIXct(DIV\_M$Date)

DIV\_M <- DIV\_M[ nrow(DIV\_M):1, ]

DIV\_M<- DIV\_M[-1:-1416,]

temp <- c("2018-01-31",1.87)

DIV\_M <- rbind(DIV\_M, temp)

DIV\_M$Value = as.numeric(DIV\_M$Value)

#

# for (i in seq\_along(t$Div\_rate\_daily)){

# if (i < 16)

# t$Div\_rate\_monthly[i] = mean(t$Div\_rate\_daily[i:(i+30)], na.rm=TRUE)

# else

# t$Div\_rate\_monthly[i] = mean(t$Div\_rate\_daily[(i-15):(i+15)], na.rm=TRUE)}

#

# ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=Div\_rate\_daily))

# ggplot(t,aes(x=Date)) +

# geom\_line(aes(y=Div\_rate\_monthly))

#

# stat.desc(t$Div\_rate\_daily)

# describe(t$Div\_rate\_daily)

# stat.desc(t$Div\_rate\_monthly)

# describe(t$Div\_rate\_monthly)

for (i in seq\_along(t$Div\_rate\_daily)){

for (j in seq\_along(DIV\_M$Value)){

if (t$Date[i] <= DIV\_M$Date[j+1] && t$Date[i] > DIV\_M$Date[j]){

t$Div\_rate\_monthly[i] = DIV\_M$Value[j+1]/100

break

}

}

}

```

#save to tt

```{r }

tt <- t

```

#loop machine learning

```{r }

table\_machine <- data.frame(i=as.numeric(),j=as.numeric(),port1\_mean=as.numeric(),port1\_var=as.numeric(),port1\_sharp=as.numeric(),port2\_mean=as.numeric(),port2\_var=as.numeric(),port2\_sharp=as.numeric())

count = 1

# loop

for (i in seq(10, 250, 10)){

for (j in seq(10, 250, 10)){

if (i<j){

t <- tt

t$SMA60 <- SMA(t$Close,i)

t$SMA120 <- SMA(t$Close,j)

t<- t[-1:-251,] # here remove training 120

for (k in seq\_along(t$SMA120)){

if (t$SMA60[k]>t$SMA120[k])

t$signal[k] = 1 # signal = 1 is buy singal

else if (t$SMA60[k]<t$SMA120[k])

t$signal[k] = 0

else

t$signal[k] = t$singal[k-1]

}

# port1

t$port1\_Relreturn[1] = NA

for (k in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[k]== 1){

t$port1\_Relreturn[k+1] = (t$Close[k+1]-t$Close[k]+t$Close[k]\*t$Div\_rate\_monthly[k]/365\*as.numeric(difftime(t$Date[k+1],t$Date[k]))) / t$Close[k] \* 100000}

else{

t$port1\_Relreturn[k+1] = (t$Close[k]-t$Close[k+1]) / t$Close[k] \* 100000

}

}

# port 2

t$port2\_Relreturn[1] = NA

for (k in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[k]== 1){

aa <- EuropeanOption("call", t$Close[k+1], t$Close[k],t$Div\_rate\_monthly[k+1], (t$Free\_rate[k+1]/100), (90-as.numeric(difftime(t$Date[k+1],t$Date[k])))/365, (t$VIX[k+1]/100))

bb <- EuropeanOption("call", t$Close[k], t$Close[k],t$Div\_rate\_monthly[k], (t$Free\_rate[k]/100), 0.246575342, (t$VIX[k]/100))

} else {

aa <- EuropeanOption("put", t$Close[k+1], t$Close[k],t$Div\_rate\_monthly[k+1], (t$Free\_rate[k+1]/100), (90-as.numeric(difftime(t$Date[k+1],t$Date[k])))/365, (t$VIX[k+1]/100))

bb <- EuropeanOption("put", t$Close[k], t$Close[k],t$Div\_rate\_monthly[k], (t$Free\_rate[k]/100), 0.246575342, (t$VIX[k]/100))

}

t$port2\_Relreturn[k+1] = (aa$value-bb$value) / bb$value \* 100000

}

t <- t[-1,]

# store data

table\_machine[count,1] = i

table\_machine[count,2] = j

table\_machine[count,3] = mean(t[,10])\*252

table\_machine[count,4] = var(t[,10])\*252

a <- as.xts(t[,10],t$Date)

b <- as.xts(t[,3]/100,t$Date)

table\_machine[count,5] = SharpeRatio(a,b,FUN = "StdDev")

# table\_machine[count,5] = mean(sharpe(t[,10],r=t[,3]/100))

table\_machine[count,6] = mean(t[,11])\*252

table\_machine[count,7] = var(t[,11])\*252

c <- as.xts(t[,11],t$Date)

table\_machine[count,8] = SharpeRatio(c,b,FUN = "StdDev")

# table\_machine[count,8] = mean(sharpe(t[,11],r=t[,3]/100))

count = count +1

}

}

}

```

##Opitimize

#Signal

```{r }

table\_machine <- read\_csv("C:/Users/mingw/Dropbox/[IAQF\_Max]/Data/table\_machine.csv")

attach(table\_machine)

head(table\_machine[order(-port2\_mean),], n=20)

head(table\_machine[order(-port2\_sharp),], n=20)

plot(table\_machine$port1\_mean,table\_machine$port2\_mean)

plot(table\_machine$port1\_sharp,table\_machine$port2\_sharp)

fff <- head(table\_machine[order(-port2\_sharp,-port2\_mean),],n=50)

alpha = .1

fff$value\_.1 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .2

fff$value\_.2 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .3

fff$value\_.3 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .4

fff$value\_.4 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

attach(fff)

head(fff[order(-port2\_sharp),], n=10)

head(fff[order(-value\_.1),], n=10)

head(fff[order(-value\_.2),], n=10)

head(fff[order(-value\_.3),], n=10)

head(fff[order(-value\_.4),], n=10)

```

#T

```{r }

t <- tt

t$SMA210 <- SMA(t$Close,210)

t$SMA220 <- SMA(t$Close,220)

t<- t[-1:-251,] # here remove training 120

for (k in seq\_along(t$SMA220)){

if (t$SMA210[k]>t$SMA220[k])

t$signal[k] = 1 # signal = 1 is buy singal

else if (t$SMA210[k]<t$SMA220[k])

t$signal[k] = 0

else

t$signal[k] = t$singal[k-1]

}

table\_machine\_T <- data.frame(T=as.numeric(),port1\_mean=as.numeric(),port1\_var=as.numeric(),port1\_sharp=as.numeric(),port2\_mean=as.numeric(),port2\_var=as.numeric(),port2\_sharp=as.numeric())

count = 1

for (T in seq(10, 370, 30)){

# port1

t$port1\_Relreturn[1] = NA

for (k in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[k]== 1){

t$port1\_Relreturn[k+1] = (t$Close[k+1]-t$Close[k]+t$Close[k]\*t$Div\_rate\_monthly[k]/365\*as.numeric(difftime(t$Date[k+1],t$Date[k]))) / t$Close[k] \* 100000}

else{

t$port1\_Relreturn[k+1] = (t$Close[k]-t$Close[k+1]) / t$Close[k] \* 100000

}

}

# port 2

t$port2\_Relreturn[1] = NA

for (k in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[k]== 1){

aa <- EuropeanOption("call", t$Close[k+1], t$Close[k],t$Div\_rate\_monthly[k+1], (t$Free\_rate[k+1]/100), (T-as.numeric(difftime(t$Date[k+1],t$Date[k])))/365, (t$VIX[k+1]/100))

bb <- EuropeanOption("call", t$Close[k], t$Close[k],t$Div\_rate\_monthly[k], (t$Free\_rate[k]/100), T/365, (t$VIX[k]/100))} else {

aa <- EuropeanOption("put", t$Close[k+1], t$Close[k],t$Div\_rate\_monthly[k+1], (t$Free\_rate[k+1]/100), (T-as.numeric(difftime(t$Date[k+1],t$Date[k])))/365, (t$VIX[k+1]/100))

bb <- EuropeanOption("put", t$Close[k], t$Close[k],t$Div\_rate\_monthly[k], (t$Free\_rate[k]/100), T/365, (t$VIX[k]/100))

}

t$port2\_Relreturn[k+1] = (aa$value-bb$value) / bb$value \* 100000

}

t <- t[-1,]

# store data

table\_machine\_T[count,1] = T

table\_machine\_T[count,2] = mean(t[,10])\*252

table\_machine\_T[count,3] = var(t[,10])\*252

a <- as.xts(t[,10],t$Date)

b <- as.xts(t[,3]/100,t$Date)

table\_machine\_T[count,4] = SharpeRatio(a,b,FUN = "StdDev")

# table\_machine\_T[count,4] = mean(sharpe(t[,10],r=t[,3]/100))

table\_machine\_T[count,5] = mean(t[,11])\*252

table\_machine\_T[count,6] = var(t[,11])\*252

c <- as.xts(t[,11],t$Date)

table\_machine\_T[count,7] = SharpeRatio(c,b,FUN = "StdDev")

# table\_machine\_T[count,7] = mean(sharpe(t[,11],r=t[,3]/100))

count = count +1

}

```

#T

```{r }

table\_machine\_T <- read\_csv("C:/Users/mingw/Dropbox/[IAQF\_Max]/Data/table\_machine\_T.csv")

#attach(table\_machine\_T)

#

# head(table\_machine[order(-port2\_mean),], n=20)

#

# head(table\_machine[order(-port2\_sharp),], n=20)

#

# plot(table\_machine$port1\_mean,table\_machine$port2\_mean)

#

# plot(table\_machine$port1\_sharp,table\_machine$port2\_sharp)

#

# fff <- head(table\_machine[order(-port2\_sharp,-port2\_mean),],n=10)

fff <- table\_machine\_T[-c(1,2,3),]

alpha = .1

fff$value\_.1 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .2

fff$value\_.2 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .3

fff$value\_.3 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .4

fff$value\_.4 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

alpha = .5

fff$value\_.5 <- fff$port2\_mean/mean(fff$port2\_mean)\*alpha + fff$port2\_sharp/mean(fff$port2\_sharp)\*(1-alpha)

attach(fff)

head(fff[order(-port2\_sharp),], n=10)

head(fff[order(-value\_.1),], n=10)

head(fff[order(-value\_.2),], n=10)

head(fff[order(-value\_.3),], n=10)

head(fff[order(-value\_.4),], n=10)

head(fff[order(-value\_.5),], n=10)

```

**Optimal\_portfolio**

---

title: "Optimal\_port2"

author: "Panther"

date: "February 19, 2018"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

##Load Data

```{r Load Data, message=FALSE}

library(quanteda); library(tidyverse); library(TTR);library(ggplot2);library(scales); library(fOptions); library(pastecs); library(psych);library(crayon);library(PerformanceAnalytics);library(stats);library(RQuantLib);

# here data read is being processed

setwd("F:/Dropbox/[IAQF\_Max]/R")

#t <- read\_csv("../data/data\_div.csv")

#t <- read\_csv("../Data/data.csv")

#t <- read\_csv("F:/Dropbox/[IAQF]/Data/data.csv")

# t$Date <- strptime(as.character(t$Date), "%m/%d/%Y")

# format(t$Date, "%Y-%m-%d")

# t$Date <- as.Date(t$Date)

```

#Process Data

```{r Process Data}

GSPC <- read\_csv("../data/^GSPC.csv")

Treasury <- read\_csv("../data/DTB3.csv")

VIX <- read\_csv("../data/^VIX.csv")

DIV <- read\_csv("../data/Dividends.csv")

DIV\_M <- read\_csv("../data/SP500\_DIV\_YIELD\_MONTH.csv")

# DIV$`SPX Index` <- strptime(as.character(DIV$`SPX Index`), "%m/%d/%Y")

# format(DIV$`SPX Index`, "%Y-%m-%d")

# DIV$`SPX Index` <- as.Date(DIV$`SPX Index`)

# t <- t[-10]

t<-cbind(GSPC,Treasury$DTB3[match(GSPC$Date,Treasury$DATE)])

t<-cbind(t,VIX$Close[match(t$Date,VIX$Date)])

t<-cbind(t,DIV$X2[match(t$Date,DIV$`SPX Index`)])

t<-t[,c(1,5,8,9,10)]

t<- t[-1:(-10054+250),] # extra 120 for training signal, will remove later

colnames(t)[3] <- "Free\_rate"

colnames(t)[4] <- "VIX"

colnames(t)[5] <- "Div\_rate\_daily"

t$Free\_rate = as.numeric(as.character(t$Free\_rate))

t$Div\_rate\_daily = as.numeric(as.character(t$Div\_rate\_daily))

```

#Missing Data

```{r Missing Data}

#missing data of free rate

temp <- which(is.na(t$Free\_rate))

for (i in temp)

t$Free\_rate[i] <- mean(t$Free\_rate[(i-5):(i+5)], na.rm=TRUE)

temp <- which(is.na(t$Div\_rate\_daily))

for (i in temp){

if (i<15)

t$Div\_rate\_daily[i] <- mean(t$Div\_rate\_daily[1:30], na.rm=TRUE)

else

t$Div\_rate\_daily[i] <- mean(t$Div\_rate\_daily[(i-15):(i+15)], na.rm=TRUE)}

```

#Smooth dividend daily rate as monthly rate

```{r smooth dividend}

t$Div\_rate\_monthly <- t$Div\_rate\_daily

# t$POSIXct\_Date <- as.POSIXct(t$Date)

# DIV\_M$POSIXct\_Date <- as.POSIXct(DIV\_M$Date)

DIV\_M <- DIV\_M[ nrow(DIV\_M):1, ]

DIV\_M<- DIV\_M[-1:-1416,]

temp <- c("2018-01-31",1.87)

DIV\_M <- rbind(DIV\_M, temp)

DIV\_M$Value = as.numeric(DIV\_M$Value)

#

for (i in seq\_along(t$Div\_rate\_daily)){

for (j in seq\_along(DIV\_M$Value)){

if (t$Date[i] <= DIV\_M$Date[j+1] && t$Date[i] > DIV\_M$Date[j]){

t$Div\_rate\_monthly[i] = DIV\_M$Value[j+1]/100

break

}

}

}

```

#Moving Average and momentum signal

```{r momentum signal}

t$SMA210 <- SMA(t$Close,210)

t$SMA220 <- SMA(t$Close,220)

t<- t[-1:-250,] # here remove training 120

for (k in seq\_along(t$SMA220)){

if (t$SMA210[k]>t$SMA220[k])

t$signal[k] = 1 # signal = 1 is buy singal

else if (t$SMA210[k]<t$SMA220[k])

t$signal[k] = 0

else

t$signal[k] = t$singal[k-1]

}

```

## Create portfolios

Absolute Return = shift between one lag of close prices

Percentage Return = Absolute Return / last day close price

Relative Return = Percentage return \* initiate price($100,000)

We use Relative Return as analysis object

#portfolio 1

Buy/short 1 index on today's closing and close position next day

Assume when you short today, you buy an index for paying back next day

Recieve dividend while buying

```{r portfolio 1}

t$port1\_absreturn[1] = NA

#t$port1\_Perreturn[1] = NA

t$port1\_Relreturn[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

t$port1\_absreturn[i+1] = (t$Close[i+1]-t$Close[i]+t$Close[i]\*t$Div\_rate\_monthly[i]/365\*as.numeric(difftime(t$Date[i+1],t$Date[i])))

# t$port1\_Perreturn[i+1] = (t$Close[i+1]-t$Close[i]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i+1]-t$Close[i]+t$Close[i]\*t$Div\_rate\_monthly[i]/365\*as.numeric(difftime(t$Date[i+1],t$Date[i]))) / t$Close[i] \* 100000

}

else{

t$port1\_absreturn[i+1] = t$Close[i]-t$Close[i+1]

# t$port1\_Perreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i]

t$port1\_Relreturn[i+1] = (t$Close[i]-t$Close[i+1]) / t$Close[i] \* 100000

}

}

t$port1\_Perreturn <- t$port1\_Relreturn/100000

```

#portfolio 2

Buy at the money call/put option(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 2}

t$port2\_absreturn[1] = NA

#t$port2\_Perreturn[1] = NA

t$port2\_Relreturn[1] = NA

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

if (t$signal[i]== 1){

aa <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (365-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("call", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 1, (t$VIX[i]/100))

} else {

aa <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (365-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb <- EuropeanOption("put", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 1, (t$VIX[i]/100))

}

t$port2\_absreturn[i+1] = aa$value-bb$value

# t$port2\_Perreturn[i+1] = (aa$value-bb$value) / bb$value

t$port2\_Relreturn[i+1] = (aa$value-bb$value) / bb$value \* 100000

}

t$port2\_Perreturn <- t$port2\_Relreturn/100000

```

#portfolio 3

Buy at the money straddle(90 days maturity) on today's closing and close position next day

Volatility is VIX

Risk free rate is 90 days T-bill rate

Consider dividends

```{r portfolio 3}

t$port3\_absreturn[1] = NA

#t$port3\_Perreturn[1] = NA

t$port3\_Relreturn[1] = NA

#ca <- GBSOption("c",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cb <- GBSOption("p",t$Close[1],t$Close[1],0.246575342,(t$Free\_rate[1]/100),0,(t$VIX[1]/100))

#cc = ca@price + cb@price

for (i in seq\_along(t$Close[1:(length(t$Close)-1)])){

aa1 <- EuropeanOption("call", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (365-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

aa2 <- EuropeanOption("put", t$Close[i+1], t$Close[i],t$Div\_rate\_monthly[i+1], (t$Free\_rate[i+1]/100), (365-as.numeric(difftime(t$Date[i+1],t$Date[i])))/365, (t$VIX[i+1]/100))

bb1 <- EuropeanOption("call", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 1, (t$VIX[i]/100))

bb2 <- EuropeanOption("put", t$Close[i], t$Close[i],t$Div\_rate\_monthly[i], (t$Free\_rate[i]/100), 1, (t$VIX[i]/100))

t$port3\_absreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value)

# t$port3\_Perreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)

t$port3\_Relreturn[i+1] = (aa1$value+aa2$value-bb1$value-bb2$value) / (bb1$value+bb2$value)\*100000

}

t$port3\_Perreturn <- t$port3\_Relreturn/100000

```

#Dump first day NA return

```{r}

t <- t[-1,]

```

#Momentum effect of portfolio returns

Obvious autocorrelation, large ARMA parameters

```{r Momentum effect}

acf(t$port1\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port1\_Relreturn,lag.max = 50)

acf(t$port2\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port2\_Relreturn,lag.max = 50)

acf(t$port3\_Relreturn,lag.max = 20,type='correlation')

pacf(t$port3\_Relreturn,lag.max = 50)

```

# Scatter of RelReturn

```{r Scatter Plot RelReturn}

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port1\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port2\_Relreturn),size =1, shape= 1)

ggplot(t,aes(x=Date)) +

geom\_point(aes(y=port3\_Relreturn),size =1, shape= 1)

```

#Statistics Comparison

```{r Statistics table}

base <- tt[,c(11,13,15)]

#base <- t[,c(11,14,17)]

colnames(base) <- c("port1\_Relreturn","port2\_Relreturn","port3\_Relreturn")

a1 <- as.xts(base[,1],t$Date)

a2 <- as.xts(base[,2],t$Date)

a3 <- as.xts(base[,3],t$Date)

b <- as.xts(t[,3]/100,t$Date)

c1 = SharpeRatio(a1,b,FUN = "StdDev")

c2 = SharpeRatio(a2,b,FUN = "StdDev")

c3 = SharpeRatio(a3,b,FUN = "StdDev")

stat <- sapply(base,sum) %>%

rbind(sapply(base,mean)) %>%

rbind(sapply(base,mean)\*252) %>%

rbind(sapply(base,var)) %>%

rbind(sapply(base,skew)) %>%

rbind(sapply(base,kurtosis)) %>%

rbind(c(c1,c2,c3))

rownames(stat) <- c("sum","mean","annual mean","var","skew","kurtosis","sharp ratio")

print(stat)

```

#Cumulative RelReturn

See how the cost of options and straddles ruined your return

```{r Cumulative Relreturn}

t <- read\_csv("../data/data\_optimal.csv")

tt <- read\_csv("../data/data\_final.csv")

t <- t[-1,]

tt <- tt[-1,]

#t$port1\_cum\_Relreturn <- cumsum(t$port1\_Relreturn)

t$port2\_cum\_Relreturn <- cumsum(t$port2\_Relreturn)

#t$port3\_cum\_Relreturn <- cumsum(t$port3\_Relreturn)

t$initial <- cumsum(tt$port2\_Relreturn)

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_cum\_Relreturn,color="Optimal")) +

geom\_line(aes(y=initial,color="Initial")) +

theme\_bw() +

ggtitle("Cumulative Return comparison")

```

#Realized Volatility on Relreturn

Realized Vol is limited, 10 years has upper boundary

```{r Realized Volatility}

t$port1\_real\_vol <- 252\*cumsum(t$port1\_Relreturn^2)/seq(length(t$port1\_Relreturn)) %>% sqrt()

t$port2\_real\_vol <- 252\*cumsum(t$port2\_Relreturn^2)/seq(length(t$port2\_Relreturn)) %>% sqrt()

t$port3\_real\_vol <- 252\*cumsum(t$port3\_Relreturn^2)/seq(length(t$port3\_Relreturn)) %>% sqrt()

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_real\_vol,color="Port1")) +

geom\_line(aes(y=port2\_real\_vol,color="Port2")) +

geom\_line(aes(y=port3\_real\_vol,color="Port3")) +

theme\_bw() +

ggtitle("Realized Volatility in all time")

```

#Rolling Analysis annually(252 days)

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis annually}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,252,gap=252,by=1,FUN="mean")

t$port2\_roll\_mean <- apply.rolling(port2,252,gap=252,by=1,FUN="mean")

t$port3\_roll\_mean <- apply.rolling(port3,252,gap=252,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,252,gap=252,by=1,FUN="sd")

t$port2\_roll\_vol <- apply.rolling(port2,252,gap=252,by=1,FUN="sd")

t$port3\_roll\_vol <- apply.rolling(port3,252,gap=252,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,252)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,252)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,252)

```

#Rolling Analysis 60 days

Rolling mean is smoother, bur going worse.

Rolling vlo is controlled.

Rolling correlation is towards positive, which is not we expecting.

```{r Rolling Analysis 60 days}

port1 <- xts(x=t$port1\_Relreturn, order.by=t$Date)

t$port1\_roll\_mean <- apply.rolling(port1,60,gap=60,by=1,FUN="mean")

t$port1\_roll\_vol <- apply.rolling(port1,60,gap=60,by=1,FUN="sd")

port2 <- xts(x=t$port2\_Relreturn, order.by=t$Date)

t$port2\_roll\_mean <- apply.rolling(port2,60,gap=60,by=1,FUN="mean")

t$port2\_roll\_vol <- apply.rolling(port2,60,gap=60,by=1,FUN="sd")

port3 <- xts(x=t$port3\_Relreturn, order.by=t$Date)

t$port3\_roll\_mean <- apply.rolling(port3,60,gap=60,by=1,FUN="mean")

t$port3\_roll\_vol <- apply.rolling(port3,60,gap=60,by=1,FUN="sd")

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_mean,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port1\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port2\_roll\_vol,na.rm = TRUE))

ggplot(t,aes(x=Date)) +

geom\_line(aes(y=port3\_roll\_vol,na.rm = TRUE))

chart.RollingCorrelation(t$port1\_roll\_mean, t$port1\_roll\_vol,60)

chart.RollingCorrelation(t$port2\_roll\_mean, t$port2\_roll\_vol,60)

chart.RollingCorrelation(t$port3\_roll\_mean, t$port3\_roll\_vol,60)

```