

Condition_Research

Jakub Skrajny

11 01 2021

Introduction

This is my project for WNE UW course called “Program R w zastosowaniach ekonomicznych i finansowych”. My goal was to analyze how choice of a norm influence the condition number of matrices.

Dataset

In order to gain data for this project I created a script.

This script:

- generates some square matrices using normal distribution
- counts condition numbers of matrices based on different norms
- saves it to csv (size of data that I generated makes it unnecessary actually)

Theses are norms that i used:

- 2-norm
- Frobenius norm
- max norm
- infinity norm

```
# number of matrices
K<-1000

# size of matrix
N<-10

# normal distribution parameters
sd<-100
mean<-0

# containers for conditions
norm2_cond <- numeric(0)
normF_cond <- numeric(0)
normM_cond <- numeric(0)
normI_cond <- numeric(0)

# main
for (i in 1:K) {
  A <- matrix( rnorm(N*N, mean, sd), N, N)
```

```

AI <- solve(A)
norm2_cond <- append(norm2_cond, norm(A, '2')*norm(AI, '2'))
normF_cond <- append(normF_cond, norm(A, 'F')*norm(AI, 'F'))
normM_cond <- append(normM_cond, norm(A, 'M')*norm(AI, 'M'))
normI_cond <- append(normI_cond, norm(A, 'I')*norm(AI, 'I'))
}

df<-data.frame(norm2_cond, normF_cond, normM_cond, normI_cond)
write.csv(df, paste(getwd(), "/conditions.csv", sep=""), row.names = FALSE)

head(df)

```

```

##   norm2_cond normF_cond normM_cond normI_cond
## 1   29.67608   55.85232   5.714873   93.37723
## 2  331.82828  618.69322  44.363943  804.03083
## 3   62.32005  109.14510  11.594216  208.48294
## 4   12.41014   29.99160   2.065689   45.31688
## 5   16.02342   32.65721   2.292479   65.77552
## 6   10.10517   27.99624   1.598895   32.13902

```

Summary table

```
summary(df)
```

##	norm2_cond	normF_cond	normM_cond	normI_cond
## Min. :	6.138	Min. : 18.46	Min. : 1.124	Min. : 20.64
## 1st Qu.:	18.907	1st Qu.: 39.48	1st Qu.: 3.497	1st Qu.: 60.46
## Median :	30.930	Median : 60.01	Median : 5.653	Median : 99.25
## Mean :	155.838	Mean : 286.61	Mean : 27.948	Mean : 511.01
## 3rd Qu.:	64.388	3rd Qu.: 116.48	3rd Qu.: 11.600	3rd Qu.: 205.21
## Max. :	22742.863	Max. : 43754.04	Max. : 4594.242	Max. : 88917.81

We see that the largest values are extremely different than others.
It means that even if we could properly count most of the equations with matrices.
There are some matrices that we probably shouldn't even try to use.

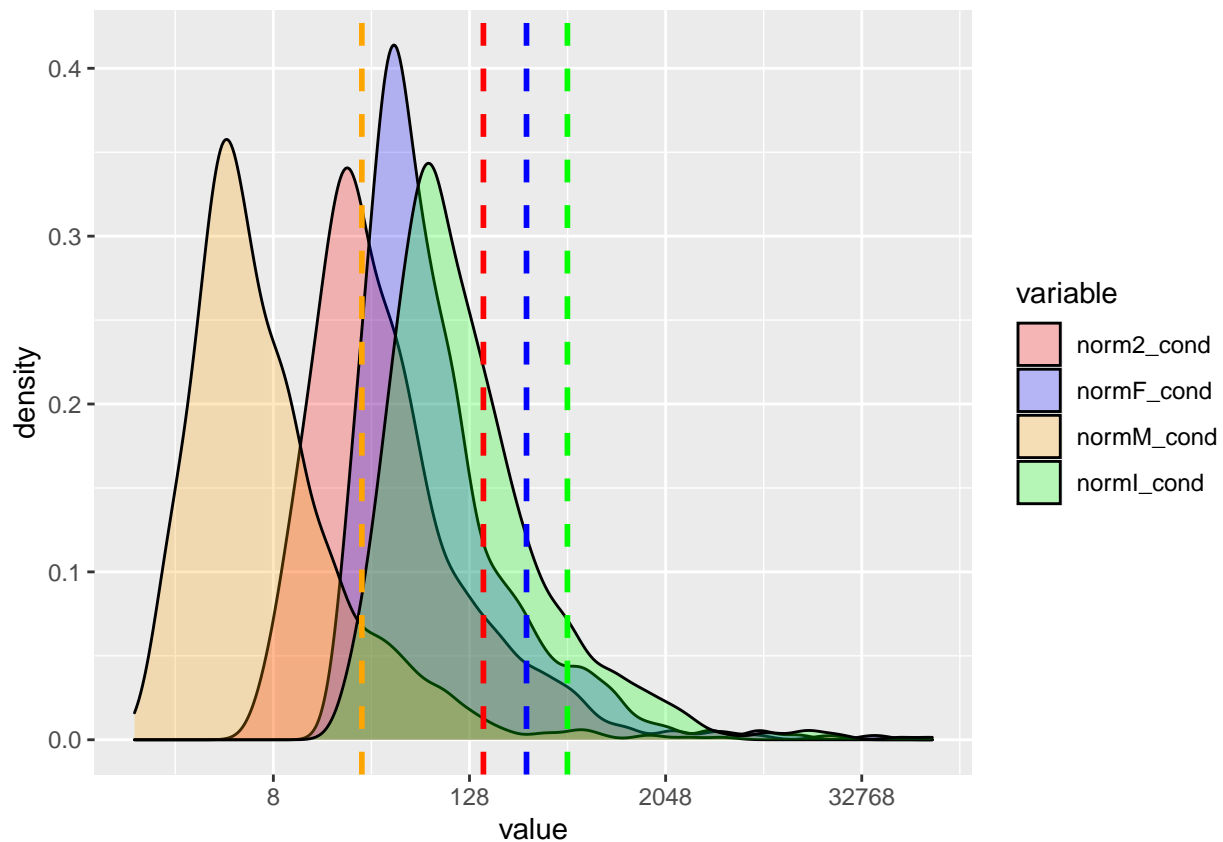
Density plots

This is a diagram that shows density distributions that are made from conditional numbers.
As we expected, there are a lot of relatively small values.
Shapes are very similar to each other.
We can suppose that correlation between these functions is very high.

```
melted <- melt(df)
```

```
## No id variables; using all as measure variables
```

```
ggplot(melted, aes(x=value, fill=variable)) +
  geom_density(alpha=0.25) +
  scale_fill_manual(values = c("red", "blue", "orange", "green")) +
  geom_vline(aes(xintercept=mean(norm2_cond)), color="red", linetype="dashed", size=1) +
  geom_vline(aes(xintercept=mean(normF_cond)), color="blue", linetype="dashed", size=1) +
  geom_vline(aes(xintercept=mean(normM_cond)), color="orange", linetype="dashed", size=1) +
  geom_vline(aes(xintercept=mean(normI_cond)), color="green", linetype="dashed", size=1) +
  scale_x_continuous(trans='log2')
```

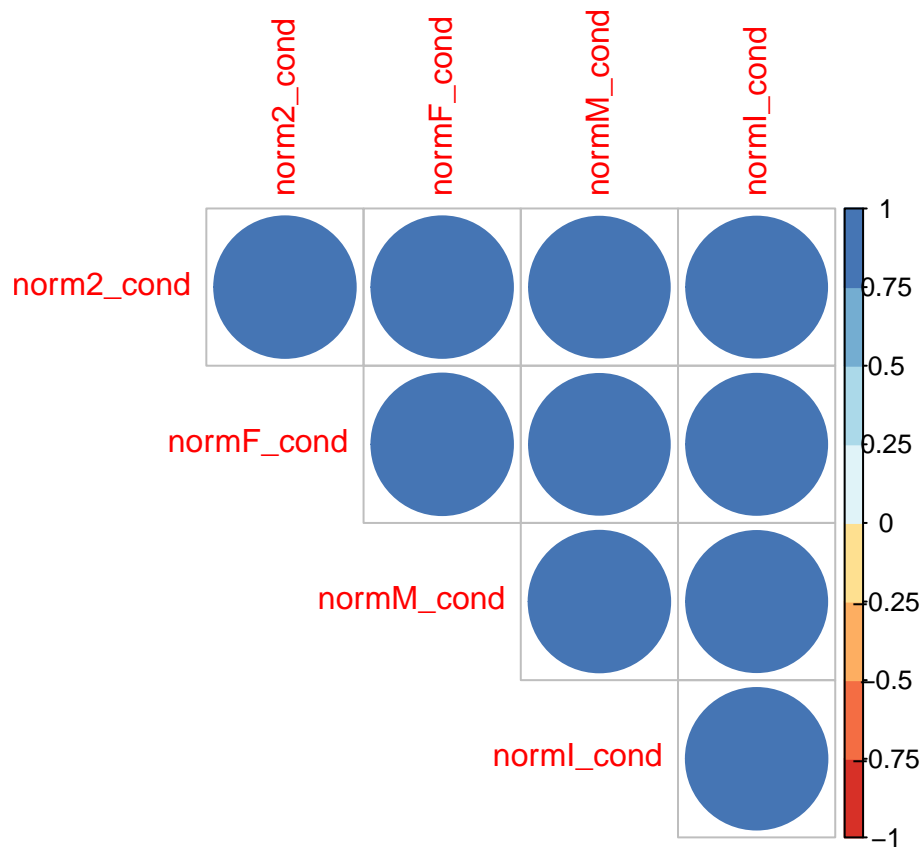


Correlations between condition numbers based on different norms

As we thought, these functions correlate with each other totally.

This might mean that no matter what norm we choose to check condition number, we get similar information about difficulty that this matrix can cause during calculations.

```
C <-cor(df)
corrplot(C, type="upper", order="hclust",
  col=brewer.pal(n=8, name="RdYlBu"))
```



Summary

If it comes to matrices created using normal distribution, there is no big difference what norm we choose to count condition number.