# NS-Modelle

|  |  |
| --- | --- |
|  |  |
| vorgelegt am | 08.03.2022 |
| von cand. | Lior Rosilio, Matrikel-Nr. 11116604  Thu Ha Tran, Matrikel-Nr. 11134861 |
| E-Mail-Adressen | [lioradler91@gmail.com](mailto:lioradler91@gmail.com)  thuhatran0298@gmail.com |

Inhaltsverzeichnis

AbbildungsverzeichnisI

AbkürzungsverzeichnisII

Aufgabe 11

Aufgabe 22

Aufgabe 37

Schlussfolgerung8

LiteraturverzeichnisIII

AnhangIV

Abbildungsverzeichnis

[Abbildung 1: Anzahl der Anleihen gemäß ihren Laufzeiten zu einem beliebigen Zeitpunkt 7](#_Toc75892950)

[Abbildung 2: Geschätzte Zinsstrukturkurve im Vergleich zur empirischen Zinsstrukturkurve 8](#_Toc75892950)

Abkürzungsverzeichnis

MSE Mean of Squared Errors

OLS Ordinary Least Square

YTM Yield to Maturity

**1. Stellen Sie das Modell von Nelson und Siegel (1987) zur Schätzung der Spot-Rate-Zinskurve dar und erläutern Sie das in dem Modell vorhandene Problem der Multikollinearität.**

Das Nelson-Siegel-Modell ist das bekannteste Instrument zur Anpassung der Renditekurve, wenn es um die Schätzung und Darstellung der Laufzeitstruktur von Zinssätzen geht. Nach Litterman & Scheinkman (1991) lässt sich der größte Teil der Renditeschwankungen bei allen festverzinslichen Wertpapieren, die durch die Zinsstrukturkurve dargestellt werden, durch drei Eigenschaften der Zinsstrukturkurve erklären: Höhe, Steigung und Krümmung. Das Nelson-Siegel-Modell verwendet einen einzigen Satz von Parametern zur Charakterisierung der gesamten Forward-Kurve, der das langfristige Niveau der Zinssätze, die Neigung der Kurve und die Buckel in der Kurve umfasst.

Die von Nelson und Siegel 1987 eingeführte Spot-Rate-Funktion zum Zeitpunkt der Fälligkeit wird wie folgt spezifiziert:

|  |  |
| --- | --- |
|  | (1) |

Die Kurve des Spot-Rates ist eine Exponentialfunktion mit drei Komponenten. Die erste Komponente, , ist eine Konstante, die ein langfristiges Niveau von Nullzinsen darstellt. Der Spot-Rate nähert sich an , wenn die Zeit bis zur Fälligkeit gegen unendlich tendiert und der Beitrag der beiden anderen Terme verschwindet. Die zweite Komponente, , steht für einen exponentiellen Zeitverfall und bestimmt die Steigung der Zinsstrukturkurve. Sie ermöglicht es der Zinsstrukturkurve, sich nach oben zu neigen, wenn ist, oder sich nach unten zu neigen, wenn ist. Diese Komponente nähert sich wiederum 0, wenn die Zeit bis zur Fälligkeit gegen unendlich tendiert, und nähert sich , wenn die Zeit bis zur Fälligkeit gegen 0 tendiert. Anders ausgedrückt: Je kürzer die Zeit bis zur Fälligkeit, desto wahrscheinlicher ist es, dass der Spot-Rate entspricht. Die dritte Komponente schließlich erzeugt entweder einen Buckel (wenn )) oder einen Tiefpunkt (wenn ) in der Zinskurve. Er beginnt bei 0, steigt an (wenn positiv ist) oder sinkt (wennnegativ ist), und nähert sich dann 0, wenn die Zeit bis zur Fälligkeit ins Unendliche wächst. Der Faktor wird verwendet, um sowohl die Form des Buckels/Tiefs als auch seine Position in der Renditekurve zu bestimmen. Die Nelson-Siegel-Gleichung ist in starkem Maße nichtlinear, da sie mehrere Variablen ( und ) enthält, und es gibt auch mehr als zwei Lösungen für diese Gleichung. Durch die Festlegung von auf einen bestimmten Wert können wir das Modell linearisieren und die Schätzung der Koeffizienten erleichtern.

Es gibt mehrere Möglichkeiten, die Parameter des NS-Modells zu schätzen. Da das Modell in hohem Maße nichtlinear ist, haben Wissenschaftler sowohl lineare als auch nichtlineare Optimierungsverfahren eingeführt. Der nichtlineare Ansatz schien jedoch zu instabilen Ergebnissen zu führen und sogar negative langfristige Raten zu erzeugen, weshalb der lineare Ansatz beliebter war. Nelson und Siegel (1987) linearisierten ihr Modell, indem sie einen festen Wert von wählten und OLS auf einem Netz von mit Werten zwischen 0,027 und 1 durchführten. Andere Wissenschaftler haben auf der Grundlage ihres Vorwissens einen festen Wert für gewählt und ebenfalls OLS durchgeführt. Wenn der Wert für für den gesamten Datensatz vorausgewählt wird, ist die Korrelation zwischen den Regressoren als trivial zu bezeichnen. Dieser Ansatz bietet jedoch kein universelles Verfahren zur Lösung des Problems der Multikollinearität.

Ein wesentliches Ziel jeder Regressionsanalyse ist es, die Beziehung zwischen den einzelnen unabhängigen Variablen und der abhängigen Variablen zu isolieren. Bei der Interpretation eines Regressionsmodells stellt der Regressionskoeffizient die durchschnittliche Veränderung der abhängigen Variablen für jede Einheitsänderung der unabhängigen Variablen dar, unter der Bedingung, dass alle unabhängigen Variablen konstant und unkorreliert gehalten werden. Wenn jedoch eine Korrelation zwischen den unabhängigen Variablen vorliegt, bedeutet dies, dass die Veränderung einer Variablen eine andere Variable beeinflussen kann. Mit anderen Worten: Wenn mehrere unabhängige Variablen stark korreliert sind, tritt Multikollinearität auf.

Multikollinearität kann zu Ungenauigkeiten bei der Schätzung der Regressionskoeffizienten führen. Das Verfahren der kleinsten Quadrate beinhaltet die Invertierung der Matrix , wobei eine Matrix ist, mit . Im Fall von perfekter Multikollinearität (die auftritt, wenn die Korrelation zwischen zwei oder mehr unabhängigen Variablen ist) ist keine Vollrangmatrix mehr und damit auch . Dies bedeutet, dass nicht invertiert werden kann. In anderen Fällen von Multikollinearität kann zwar invertiert werden, ist aber schlecht konditioniert. Daher ist die Regression sehr empfindlich, da kleine Abweichungen einen großen Einfluss auf das Ergebnis haben können.

Wissenschaftler haben berichtet, dass es eine hohe Korrelation zwischen der Steigung und der Krümmungskomponente des NS-Modells gibt. Durch Experimente mit verschiedenen Zeitreihen von Anleihen kamen Annaert et al. zu dem Schluss, dass die Korrelation zwischen der Steigung und der Krümmungskomponente stark von der Wahl des Formparameters abhängt. Sie wiesen auch darauf hin, dass der Vektor, der eine Reihe von kurzen Laufzeiten enthält, stärker durch das Problem der Kollinearität beeinträchtigt wird.

**2. Erläutern Sie in eigenen Worten den von Annaert et al. (2013) beschriebenen Ansatz zur Linderung des Multikollinearitätsproblems im Rahmen der Parameterschätzung des Nelson-Siegel-Modells. Gehen Sie dabei explizit auf die Methode der Ridge-Regression ein, indem Sie das dem Verfahren zugrundeliegende Optimierungsproblem sowie dessen Lösung darstellen und in eigenen Worten beschreiben**.

Annaert et al. führten einen alternativen Ansatz zu den gewöhnlichen kleinsten Quadraten ein, um die Auswirkungen der Multikollinearität zu minimieren. Es ist bekannt, dass Multikollinearität im Wesentlichen dann auftritt, wenn hohe Korrelationen zwischen mehr als zwei Variablen bestehen, was zu ungenauen Schätzungen der Regressionskoeffizienten führen würde.

Gegeben sei ein allgemeines lineares Regressionsmodell:

|  |  |
| --- | --- |
|  | (2) |

Y ist ein ( Vektor von Beobachtungen, X ist eine Matrix von Beobachtungen von Prädiktoren, ist der Vektor der unbekannten Regressionskoeffizienten und ist ein ( Vektor der experimentellen Fehler. Wir nehmen an, dass normalverteilt ist mit und .

Das Ziel eines jeden Regressionsproblems ist die Schätzung des Parametervektors . In der populärsten Methode: Ordinary Least Squares, wird so gewählt, dass die Summe der quadrierten Residuen minimiert wird, gegeben durch:

|  |  |
| --- | --- |
|  | (3) |

Der Schätzer der kleinsten Quadrate für ist durch eine Matrixgleichung gegeben:

|  |  |
| --- | --- |
|  | (4) |

wobei die Transponierung der Matrix ist und der Exponent die Inversion der Matrix bezeichnet. Es ist klar, dass die kleinste quadratische Schätzung den am wenigsten linear verzerrten Schätzer der Parameter ergibt. Der Schätzer mit der geringsten linearen Verzerrung ist jedoch nicht unbedingt der beste Schätzer, insbesondere dann nicht, wenn eine starke Korrelation zwischen zwei (oder mehr) Prognosevariablen festgestellt wird. Wenn die Terme der Matrixgleichung korreliert sind, neigen die Spalten der Matrix X dazu, linear abhängig zu sein, was zu einer Singulären führt. Die Kleinste-Quadrate-Schätzung reagiert daher sehr empfindlich auf minimale Änderungen in der beobachteten Response , was zu einer großen Varianz führt.

Die Idee hinter der Ridge-Regression ist, dass, indem kontrollierte Mengen an Verzerrungen im ursprünglichen Problem zugelassen werden, die Standardfehler minimiert werden, und die invertierte Matrix keine Determinante nahe Null mehr hat (was bedeutet, dass ihre Zeilen und Spalten linear abhängige Vektoren sind). Letztendlich führen die Lösungen nicht zu einer großen Varianz in den geschätzten Parametern.

Jede Zielfunktion muss einer Einschränkung unterworfen werden, um maximiert oder minimiert werden zu können. Im Fall der Ridge-Regression versucht der Ridge-Regression-Schätzer, die Ridge-Loss-Funktion zu minimieren, die wie folgt definiert ist:

|  |  |
| --- | --- |
|  | (5) |

Die Verlustfunktion besteht aus zwei Komponenten: der Summe der quadrierten Residuen und einen Penalty: . In der Ridge-Penalty fungiert als Tuning-Parameter, der die Höhe der Penalty bestimmt, die wir der Flexibilität des Regressionsmodells auferlegen wollen. Für hat der Penalty-Term keine Wirkung, die Ridge-Loss-Funktion ist einfach eine gewöhnliche Least-Square-Funktion. Für jedes trägt die Ridge-Penalty zur Verlustfunktion bei. Die Endwirkung eines Penalty besteht darin, die Regressionskoeffizienten vor einer Explosion zu bewahren, indem sie gegen Null, ihr Minimum, geschrumpft werden. Wenn k , nimmt die Flexibilität der Ridge-Regressionsanpassung ab, was zu einer geringeren Varianz, aber einer größeren Verzerrung führt.

Nehmen wir an, wir haben die Daten , wobei die Prädiktorvariablen und die Antworten sind.

Wir gehen außerdem davon aus, dass die so standardisiert sind, dass

Der Ridge-Schätzer kann wie folgt definiert werden:

|  |  |
| --- | --- |
|  | (6) |
| s.t. |  |

Diese Gleichung ist gleichbedeutend mit:

|  |  |
| --- | --- |
|  | (7) |

weil es eine Eins-zu-eins-Entsprechung zwischen den Parametern und gibt, wie Hastie, Tibshirani und Friedman (2009) betonen.

Wir schreiben die Gleichung dann in Matrixform um:

|  |  |
| --- | --- |
|  | (8) |

Multipliziert man alle Terme innerhalb der Klammern, erhält man:

|  |  |
| --- | --- |
|  | (9) |

Da als geschrieben werden kann, kann die obige Gleichung wie folgt umgeschrieben werden:

|  |  |
| --- | --- |
|  | (10) |

Wir leiten die Gleichung in Bezug auf ab:

|  |  |
| --- | --- |
|  | (11) |

Wir kommen an:

|  |  |
| --- | --- |
|  | (12) |

Dann setzen wir die Gleichung gleich 0 und lösen für :

|  |  |
| --- | --- |
|  | (13) |

Teilt man beide Seiten durch 2, erhält man:

|  |  |
| --- | --- |
|  | (14) |

Dann verschieben wir auf die linke Seite der Gleichung:

|  |  |
| --- | --- |
|  | (15) |

Da in beiden Termen auf der rechten Seite der Gleichung enthalten ist, ergibt sich die Gleichung, wenn man herauszieht:

|  |  |
| --- | --- |
|  | (16) |

wobei die Identitätsmatrix ist.

Durch Isolierung von erhalten wir die Lösung:

|  |  |
| --- | --- |
|  | (17) |

Es ist von größter Bedeutung, dass der Ridge-Parameter bestimmt wird, um den besten Koeffizienten für das Modell zu finden. Um den richtigen Wert von zu finden, sind eine Menge komplexer Algorithmen und eine hohe Rechenkapazität erforderlich. Es gibt jedoch einige "Vernunftanforderungen", die man dem Ridge-Regressionsschätzer auferlegen möchte. Diese Anforderungen führen nicht zu einer spezifischen Wahl des Penalty-Parameters, aber sie spezifizieren einen Wertebereich von sinnvollen Penalty-Parametern.

Annaert et al. verwendeten einen dreistufigen Ansatz zur Berechnung des optimalen . Zunächst führten sie eine Gittersuche auf der Grundlage der OLS-Regression durch, um die Schätzung von zu erhalten, die den geringsten mittleren quadratischen Fehler erzeugt. Der mittlere quadratische Fehler (MSE) eines Schätzers misst den Durchschnitt der Quadratfehler, d. h. die durchschnittliche quadratische Differenz zwischen dem geschätzten Wert und dem tatsächlichen Wert. sei der mittlere quadratische Fehler des Schätzers der Ridge-Regression, sei der mittlere quadratische Fehler des Schätzers der linearen Regression. Nach dem Theorem von Theobald (1974) gibt es , so dass .. Zweitens berechneten Annaert et al. die Bedingungszahl für das "optimale ". Sie waren der Meinung, dass die numerische Ungenauigkeit, die sich aus der Unkonditioniertheit von ergibt, vermieden werden sollte. Eine Matrix ist schlecht konditioniert, wenn ihre Bedingungszahl hoch ist. Die Bedingungszahl der Matrix mit sind die Eigenwerte von ist das Verhältnis ihres größten und kleinsten Eigenwerts, definiert als:

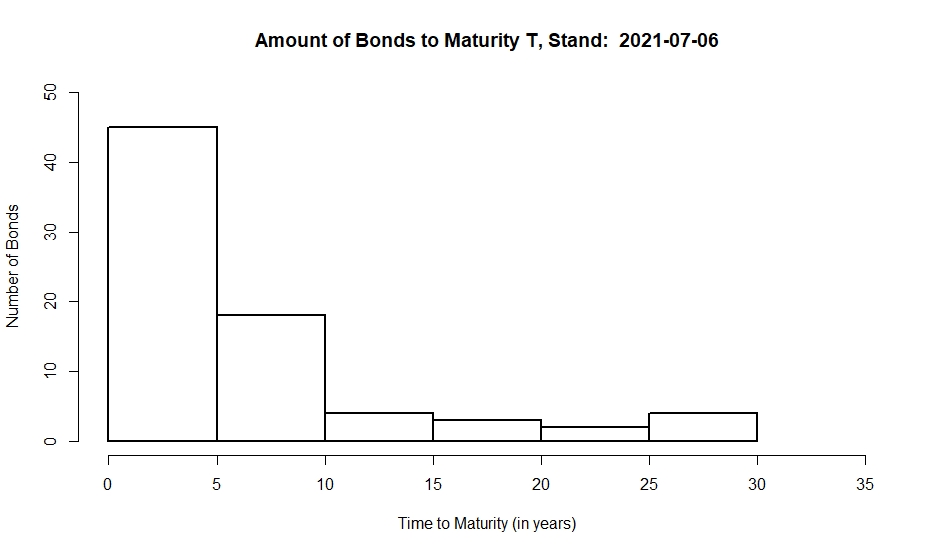
|  |  |
| --- | --- |
|  | (18) |

Wenn in den Spalten der Regressoren keine Korrelation festgestellt wird, dann ist die Matrix gut konditioniert und die Konditionszahl ist eins. Wenn eine Korrelation besteht, dann wächst die Bedingungszahl mit der Differenz zwischen dem größten und dem kleinsten Eigenwert. Hier wählten Annaert et al. eine Bedingungszahl von 10. Drittens wurden die Koeffizienten mit Hilfe der Ridge-Regression neu geschätzt, wenn die Bedingungszahl über 10 lag. Ein iteratives Suchverfahren wurde angewandt, um die niedrigste positive Zahl zu finden, wobei der Ausgangspunkt war. Die Bedingungszahl wurde iterativ neu berechnet, nachdem die Ridge-Konstante um 0,001 erhöht worden war. Wenn die neu berechnete Bedingungszahl niedriger als 10 war, wurde der Iterationsprozess beendet und der Wert von akzeptiert.

**3. Empirisches Beispiel des von Annaert et al. (2013) beschriebenen Verfahrens zur Schätzung des Nelson-Siegel-Modells**

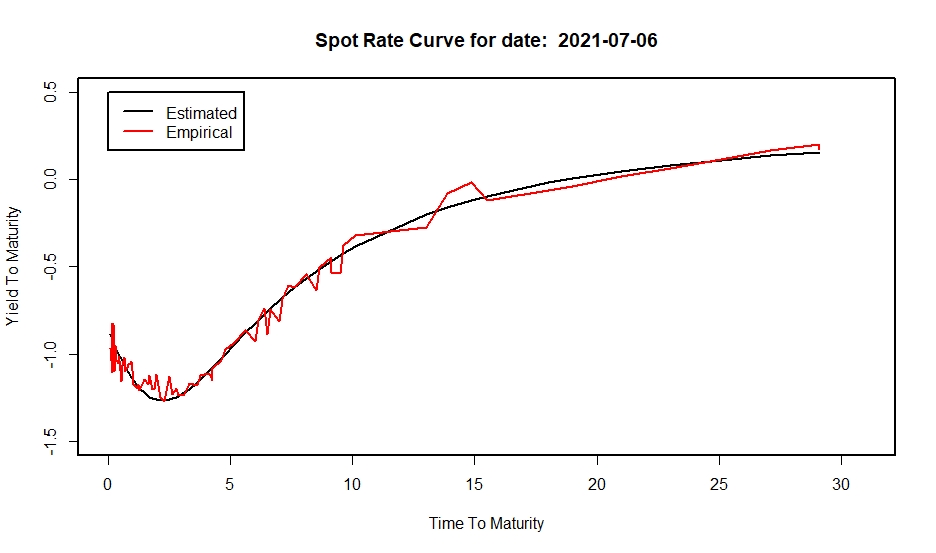
Eine Illustration des Verfahrens erfolgt im Rahmen dieser Ausarbeitung über 117 Anleihen[[1]](#footnote-1) der Bundesrepublik Deutschlands. Die Staatsanleihen und deren täglichen YTM-Änderungen im Zeitraum vom 1.1.2020 bis 31.12.2021 wurden von Eikon-Datastream gesammelt.

Der erste Schritt der Implementation des Nelson-Siegel Modells ist die Berechnung von Restlaufzeiten zu jeder täglichen Zeitperiode des Datensatzes (in Jahren). Wichtig dabei ist, dass alle nicht-emittierte Anleihen zum Zeitpunkt t nicht in die Berechnung einbezogen sind.

Abbildung 1: Anzahl der Anleihen gemäß ihren Laufzeiten zu einem beliebigen Zeitpunkt

Im nächsten Schritt führen wir die Ridge-Regression ein. Nachdem wir die optimierte lambda aus der Gleichung X anhand einer OLS-Regression bestimmt haben, ist die Konditionsnummer k aus der Gleichung X zu Berechnen. Falls die Konditionsnummer größer ist als 10, passen wir die Koeffizienten mithilfe der Gleichung X an[[2]](#footnote-2).

Schließlich erhalten wir zu jedem Zeitpunkt t eine Zinsstrukturkurve. Wichtig dabei ist, die Zinssätze zu logarithmieren.

Abbildung 2: Geschätzte Zinsstrukturkurve im Vergleich zur empirischen Zinsstrukturkurve

**4. Schlussfolgerung**

Die Grafik in der Abbildung X weist eine sehr gute Schätzung der Zinsstrukturkurve auf. Dies erfolgt auch zu unterschiedlichen Zeitpunkten wie z.B. Januar, März und September 2021. Die Multikollinearität kann durch Optimierung des Lambda-Parameters gänzlich vermieden werden, da eine Korrelation zwischen den Modellregressoren nicht vorliegt.

Literaturverzeichnis

Annaert, J., Claes, A. G., De Ceuster, M. J., & Zhang, H. (2013). Estimating the spot rate curve using the Nelson–Siegel model: A ridge regression approach. International Review of Economics & Finance, 27, 482-496.

Belsley, D. A. (1991). Conditioning diagnostics: Collinearity and weak data in regression (No. 519.536 B452). Wiley.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). An introduction to statistical learning.

Litterman, R., & Scheinkman, J. (1991). Common factors affecting bond returns. Journal of fixed income, 1(1), 54-61.

Theobald, C. M. (1974). Generalizations of mean square error applied to ridge regression. Journal of the Royal Statistical Society: Series B (Methodological), 36(1), 103-106.

Anhang

1. R-Program

# Remove previous data  
rm(list = ls())  
  
# Load string-manipulation methods  
library(stringi)

calculate\_time\_to\_maturity <- function(x,y) {  
 lubridate::time\_length(difftime(x, y), unit = "years")  
}  
  
# Read data from xlsx file  
path\_to\_xlsx = "C:/Users/lrosilio/Desktop/Deutsche Staatsanleihen.xlsx"  
data\_table = as.data.frame(readxl::read\_excel(path\_to\_xlsx))

# Use dates as row names  
rownames(data\_table) <- data\_table[,1]  
  
# Remove obsolete date column  
data\_table <- data\_table[-1]  
  
# Remove all strips and eternity bonds  
cols\_to\_remove = c()  
  
for (i in 1:ncol(data\_table)) {  
 if (grepl("strip", colnames(data\_table)[i], ignore.case = TRUE) ||  
 (is.na(stri\_extract(colnames(data\_table)[i],  
 regex = "[:digit:]{2}[:punct:][:digit:]{2}[:punct:][:digit:]{2}")))) {  
 cols\_to\_remove = c(cols\_to\_remove, i)  
 }  
}  
  
data\_table <- data\_table[-cols\_to\_remove]  
  
# Period of Validity vector  
PoV = c()  
  
for (i in 1:ncol(data\_table)) {  
 x = colnames(data\_table)[i] # Temporary variable for a better readability  
 T = stri\_extract(x, regex = "[:digit:]{2}[:punct:][:digit:]{2}[:punct:][:digit:]{2}")  
 T <- gsub("00", "20", as.Date(T, format = "%d/%m/%Y"))  
 PoV = c(PoV, T)  
 rm(x, T)  
}  
  
# Convert data into numeric type  
data\_table[] = lapply(data\_table, as.numeric)

# Calculate time to maturities  
tau = t(outer(PoV, rownames(data\_table), FUN = calculate\_time\_to\_maturity))  
tau[which(is.na(data\_table), arr.ind = TRUE)] <- NA # Omit when bond isn't traded  
  
# Convert data into matrix  
data\_table = as.matrix(data\_table)  
  
# TODO: optimize lambda  
lambda = 4  
  
# Preallocate  
b\_0 = matrix(data=NA, ncol = 1, nrow = nrow(data\_table))  
b\_1 = matrix(data=NA, ncol = 1, nrow = nrow(data\_table))  
b\_2 = matrix(data=NA, ncol = 1, nrow = nrow(data\_table))  
spot\_rates = matrix(data=NA, ncol = ncol(data\_table), nrow = nrow(data\_table))  
  
# Linear regression  
for (i in 1:nrow(data\_table)) {  
 r1 = lambda \* (1 - exp(-na.omit(tau[i,])/lambda))/na.omit(tau[i,])  
 r2 = lambda \* (1 - exp(-na.omit(tau[i,])/lambda))/na.omit(tau[i,]) - exp(-na.omit(tau[i,])/lambda)  
 fit = lm(na.omit(data\_table[i,]) ~ r1 + r2)  
   
 b\_0[i] = fit$coefficients[1]  
 b\_1[i] = fit$coefficients[2]  
 b\_2[i] = fit$coefficients[3]  
   
 # TODO: Ridge Regression  
 #olsrr::ols\_vif\_tol(fit)  
 # if VIF > 10 {  
 # re-evaluate coefficients  
 #}  
   
 spot\_rates[i,] = b\_0[i] +  
 b\_1[i] \* (lambda \* (1 - exp(-tau[i,]/lambda))/tau[i,]) +  
 b\_2[i] \* (lambda \* (1 - exp(-tau[i,]/lambda))/tau[i,] - exp(-tau[i,]/lambda))  
   
 spot\_rates[i,] = log(1 + spot\_rates[i,])  
}

# Examples - specify date for it's spot-rate curve  
date = "2021-07-06"  
x = match(as.Date(date), as.Date(rownames(data\_table)))  
  
# Histogram  
hist(tau[x,],  
 main = paste("Amount of Bonds to Maturity T, Stand: ", rownames(data\_table)[x]),  
 xlab = "Time to Maturity (in years)",  
 ylab = "Number of Bonds",  
 xlim = c(0,35),  
 ylim = c(0,50)  
)

# Make dataset continuous  
data\_table\_log = log(1 + data\_table)

# Plot  
plot(tau[x,na.omit(order(tau[x,]))],   
 spot\_rates[x,na.omit(order(tau[x,]))],  
 ylim = c(-1.5, 0.5),  
 xlim = c(0, 31),  
 main = paste("Spot Rate Curve for date: ", rownames(data\_table)[x]),  
 xlab = "Time To Maturity",  
 ylab = "Yield To Maturity",  
 type = "l")  
lines(tau[x,na.omit(order(tau[x,]))],  
 data\_table\_log[x,na.omit(order(tau[x,]))],  
 ylim = c(-1.5, 0.5),  
 xlim = c(0, 31),  
 col = "red",  
 type = "l")  
legend(0, 0.5, legend = c("Estimated", "Empirical"), col = c("black", "red"), lty=1:1)

1. zzgl. Strip Anleihen und ewige Renten [↑](#footnote-ref-1)
2. Fgl. Beasley (1991): Conditioning diagnostics – Collinearity and weak data regression [↑](#footnote-ref-2)