

Tarea. Regresión Armónica.

André Marx Puente Arévalo

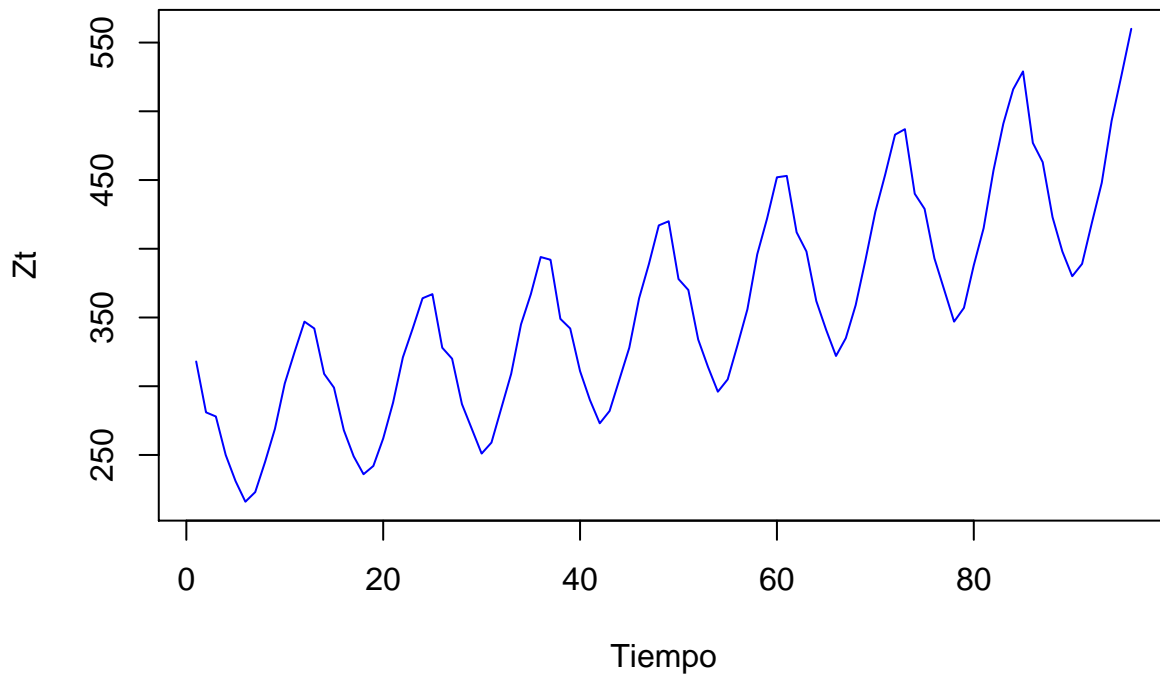
02/04/2020

```
# Leemos los datos
datos <- read.csv("/Users/andremarxpuentearevalo/Documents/Fac ciencias/Series de Tiempo/Excel/Datos RA")

# Cargamos la librería que nos permite manejar series de tiempo
library(tseries)

# Graficamos la serie para ver como se comporta
plot(datos$Zt, type = "l", col = "Blue", main = "Serie Original", ylab = "Zt", xlab = "Tiempo")
```

Serie Original

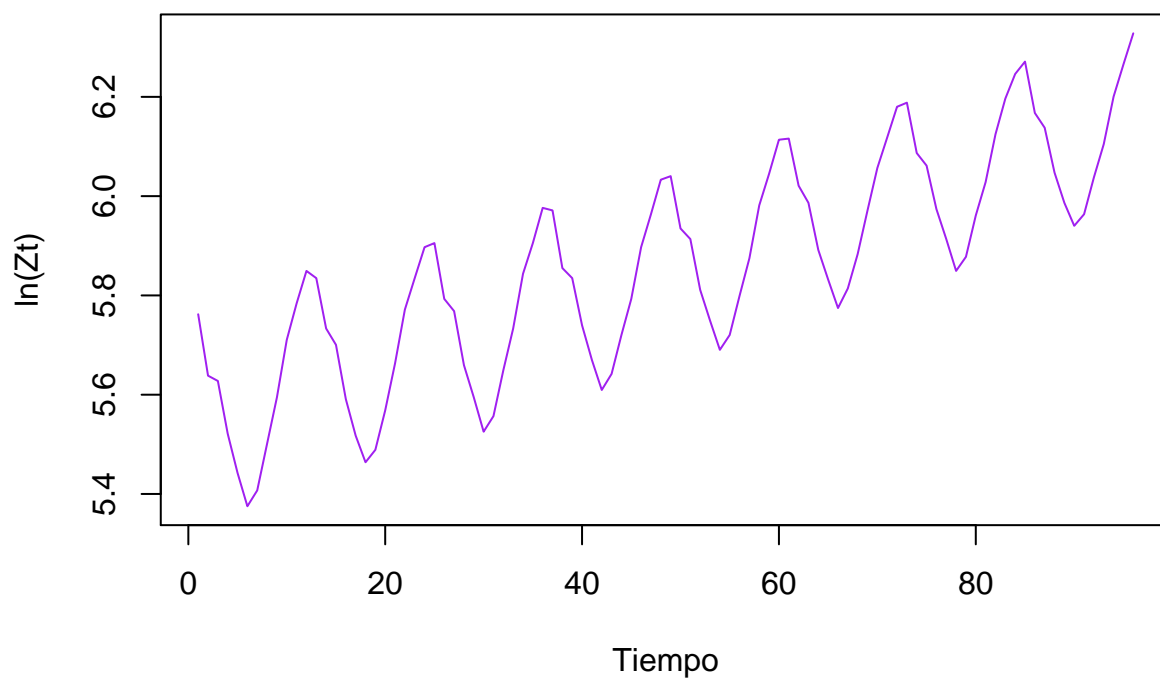


Notamos que la varianza de la serie va incrementando conforme incrementa Z_t , por lo que aplicaré una transformación para estabilizar la varianza.

```
# Aplicamos logaritmo natural a la serie
LnZt <- log(datos$Zt)

# Grafico la serie
plot(LnZt, type = "l", col = "Purple", main = "Serie transformada", ylab = "ln(Zt)", xlab = "Tiempo")
```

Serie transformada



Ahora, ya que nuestra serie tiene varianza estable, notamos que presenta tendencia, por lo que aplicaré una diferencia para quitársela. Lo que se hará es: $\nabla(\ln(Z_t))$.

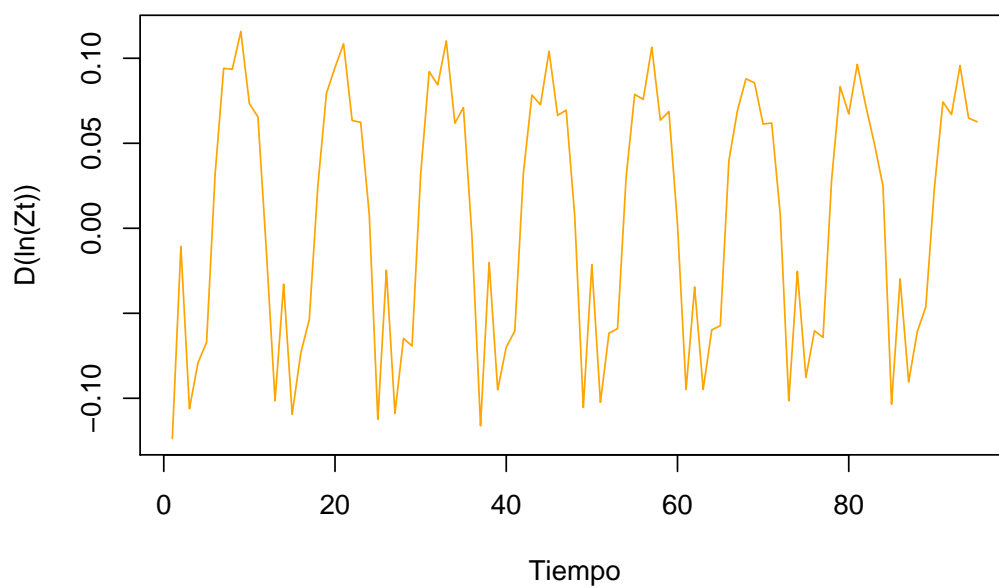
```
# Aplico la diferencia
```

```
DLnZt <- diff(LnZt)
```

```
# Grafico la serie con la primer diferencia
```

```
plot(DLnZt, type = "l", col = "Orange", main = "Serie con una diferencia", ylab = "D(ln(Zt))", xlab = "Tiempo")
```

Serie con una diferencia



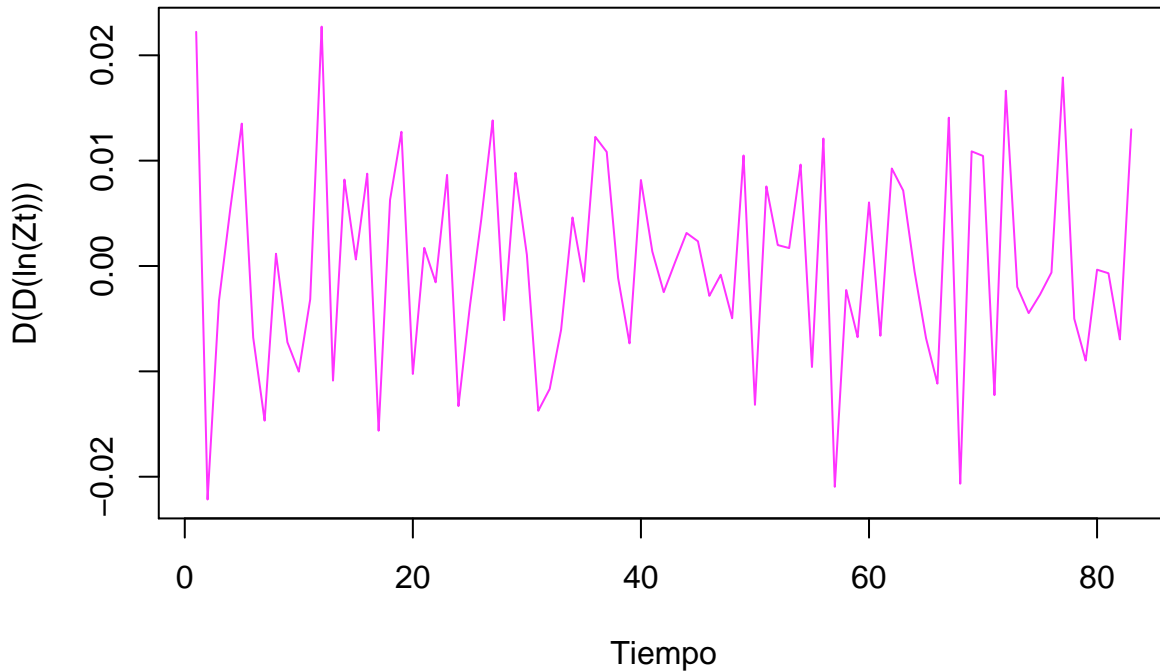
Sabemos que tiene una periodicidad de cada 12 tiempos por lo que le aplicamos una diferencia del tipo $\nabla_{12}(\nabla(\ln(Z_t)))$.

```
# Aplicamos la diferencia
DDLnZt <- diff(DLnZt,12)
```

```
# Graficamos la serie sin estacionalidad
```

```
plot(DDLnZt, type = "l", col = "#ff33ff", main = "Serie con dos diferencias", ylab = "D(D(ln(Zt)))", xlab = "Tiempo")
```

Serie con dos diferencias



Como podemos observar, ya le quité la estacionalidad a la serie, por lo que ya puedo trabajar con ella. Al final, me he quedado con 83 observaciones, pero dado que la maestra pidió que pronostiquemos las últimas 3, trabajaré con 80 observaciones, entonces tendré que crear una matriz con 78 covariables de las cuales 39 son cosenos y 39 son senos. Estos últimos se calcularán de la siguiente manera:

$$\cos(2\pi \frac{i}{n}t)$$

$$\text{sen}(2\pi \frac{i}{n}t)$$

```
# Número total de observaciones
```

```
n <- length(DDLnZt)
```

```
# Numero de senos y cosenos
```

```
m <- 39
```

```
# Inicializo una matriz
```

```
matriz <- matrix(0, nrow = n, ncol = 2*m)
```

```
# Con el siguiente for iré llenando la matriz
```

```
for (t in 1:n) {
```

```

# Para el Coseno
for (i in 1:m) {
  matriz[t, i] <- cos(2*pi*i/n*t)
}

# Para el Seno
for (j in (m+1):(2*m)) {
  matriz[t, j] <- sin(2*pi*((j-m)/n)*t)
}
}

```

Ahora, voy a armar un data frame con los senos, cosenos, sus respectivos encabezados de columna y la variable de respuesta para el entrenamiento del modelo.

```

# Vamos a crear un vector con los nombres de las columnas
nombres <- rep(0, 2*m)

for (j in 1:m) {
  nombres[j] <- gsub(" ", "", paste("C", j, "t"))
  nombres[j+m] <- gsub(" ", "", paste("S", j, "t"))
}

matriz <- as.data.frame(matriz)
names(matriz) = nombres

# Separamos la muestra en entrenamiento y en prueba
y_training <- DDLnZt[-(81:83)]

# Juntamos todos los datos
data <- cbind(as.data.frame(y_training), matriz[-(81:83), ])

head(data[,1:6], 5)

```

```

##      y_training      C1t      C2t      C3t      C4t      C5t
## 1  0.022227253  0.9971360  0.9885606  0.9743227  0.95450406  0.92921807
## 2 -0.022164148  0.9885606  0.9545041  0.8986096  0.82215601  0.72689245
## 3 -0.003296397  0.9743227  0.8986096  0.7767488  0.61499844  0.42166513
## 4  0.005509123  0.9545041  0.8221560  0.6149984  0.35188102  0.05674527
## 5  0.013518211  0.9292181  0.7268925  0.4216651  0.05674527 -0.31620767

```

A continuación, voy a aplicar el método **Backward** y el **Fordward** para la selección de mi mejor modelo. Ahora, dado que utilizaré la función “stepAIC” de la librería MASS, la selección de los modelos los hará de la siguiente manera:

- El backward lo que hace es partir de un modelo que contiene a todas las covariables y va quitando una por una hasta llegar al modelo que más minimice el AIC.
- El fordward lo que hace es partir de un modelo que no tiene covariables (sólo la constante) y va agregando una a una las covariables hasta llegar al modelo que de igual manera que el anterior, minimice más al AIC.

```

# Cargo librería
library(MASS)

# Obtenemos el modelo con todas las covariables
modeloCompleto <- lm(data$y_training ~ data$C1t +data$C2t +data$C3t +data$C4t +data$C5t +data$C6t)

```

```

# Obtenemos un modelo vacío
modeloVacio <- lm(data$y_training ~ 1, data)

# Le fijamos como una cota superior de covariables que puede tomar el metodo forward
tope <- formula(data$y_training ~ data$C1t +data$C2t +data$C3t +data$C4t +data$C5t +data$C6t

# Obtengo los modelos
modeloBackward <- stepAIC(modeloCompleto, direction = "backward", trace = F)
modeloFordward <- stepAIC(modeloVacio, scope=tope, direction = "forward", trace = F)

# Obtenemos el AIC
print(paste("El AIC del modelo completo es:", round(AIC(modeloCompleto), 3)))

## [1] "El AIC del modelo completo es: -1068.277"
print(paste("El AIC del modelo backward es:", round(AIC(modeloBackward), 3)))

## [1] "El AIC del modelo backward es: -1068.277"
print(paste("El AIC del modelo frodward es:", round(AIC(modeloFordward), 3)))

## [1] "El AIC del modelo frodward es: -763.12"
# Obtenemos el BIC
print(paste("El BIC del modelo completo es:", round(BIC(modeloCompleto), 3)))

## [1] "El BIC del modelo completo es: -877.714"
print(paste("El BIC del modelo backward es:", round(BIC(modeloBackward), 3)))

## [1] "El BIC del modelo backward es: -877.714"
print(paste("El BIC del modelo fordward es:", round(BIC(modeloFordward), 3)))

## [1] "El BIC del modelo fordward es: -617.817"
# Obtenemos el error
print(paste("La R cuadrada del modelo completo es:", round(summary(modeloCompleto)$r.squared, 5)))

## [1] "La R cuadrada del modelo completo es: 0.99987"
print(paste("La R cuadrada del modelo backward es:", round(summary(modeloBackward)$r.squared, 5)))

## [1] "La R cuadrada del modelo backward es: 0.99987"
print(paste("La R cuadrada del modelo forward es:", round(summary(modeloFordward)$r.squared, 5)))

## [1] "La R cuadrada del modelo forward es: 0.99045"

Notemos que el modelo backward tiene el mismo AIC y BIC que el modelo completo, por lo que no está
quitando ninguna covariable.

Observamos la ANOVA de dicho modelo:
modeloBackward$anova

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## data$y_training ~ data$C1t + data$C2t + data$C3t + data$C4t +

```

```
## data$C5t + data$C6t + data$C7t + data$C8t + data$C9t + data$C10t +
## data$C11t + data$C12t + data$C13t + data$C14t + data$C15t +
## data$C16t + data$C17t + data$C18t + data$C19t + data$C20t +
## data$C21t + data$C22t + data$C23t + data$C24t + data$C25t +
## data$C26t + data$C27t + data$C28t + data$C29t + data$C30t +
## data$C31t + data$C32t + data$C33t + data$C34t + data$C35t +
## data$C36t + data$C37t + data$C38t + data$C39t + data$S1t +
## data$S2t + data$S3t + data$S4t + data$S5t + data$S6t + data$S7t +
## data$S8t + data$S9t + data$S10t + data$S11t + data$S12t +
## data$S13t + data$S14t + data$S15t + data$S16t + data$S17t +
## data$S18t + data$S19t + data$S20t + data$S21t + data$S22t +
## data$S23t + data$S24t + data$S25t + data$S26t + data$S27t +
## data$S28t + data$S29t + data$S30t + data$S31t + data$S32t +
## data$S33t + data$S34t + data$S35t + data$S36t + data$S37t +
## data$S38t + data$S39t
##
## Final Model:
## data$y_training ~ data$C1t + data$C2t + data$C3t + data$C4t +
## data$C5t + data$C6t + data$C7t + data$C8t + data$C9t + data$C10t +
## data$C11t + data$C12t + data$C13t + data$C14t + data$C15t +
## data$C16t + data$C17t + data$C18t + data$C19t + data$C20t +
## data$C21t + data$C22t + data$C23t + data$C24t + data$C25t +
## data$C26t + data$C27t + data$C28t + data$C29t + data$C30t +
## data$C31t + data$C32t + data$C33t + data$C34t + data$C35t +
## data$C36t + data$C37t + data$C38t + data$C39t + data$S1t +
## data$S2t + data$S3t + data$S4t + data$S5t + data$S6t + data$S7t +
## data$S8t + data$S9t + data$S10t + data$S11t + data$S12t +
## data$S13t + data$S14t + data$S15t + data$S16t + data$S17t +
## data$S18t + data$S19t + data$S20t + data$S21t + data$S22t +
## data$S23t + data$S24t + data$S25t + data$S26t + data$S27t +
## data$S28t + data$S29t + data$S30t + data$S31t + data$S32t +
## data$S33t + data$S34t + data$S35t + data$S36t + data$S37t +
## data$S38t + data$S39t
##
##
## Step Df Deviance Resid. Df Resid. Dev AIC
## 1 1 1.006222e-06 -1297.307
```

Y efectivamente, como lo mencioné arriba, mediante el método backward no hizo ningún cambio al modelo completo. Analisemos el summary:

```
summary(modeloBackward)$coefficients
```

```
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1799185972 0.0384040825 4.6848821 0.13387898
## data$C1t 0.3578869944 0.0764781656 4.6795970 0.13402575
## data$C2t 0.3536790991 0.0754943086 4.6848445 0.13388002
## data$C3t 0.3448366315 0.0738733209 4.6679454 0.13435046
## data$C4t 0.3352085115 0.0716429506 4.6788764 0.13404579
## data$C5t 0.3217546780 0.0688412291 4.6738660 0.13418527
## data$C6t 0.3061110305 0.0655156766 4.6723326 0.13422802
## data$C7t 0.2875591550 0.0617223133 4.6589173 0.13460312
## data$C8t 0.2668162270 0.0575244955 4.6383062 0.13518347
## data$C9t 0.2455866067 0.0529916067 4.6344435 0.13529279
## data$C10t 0.2191843229 0.0481976308 4.5476161 0.13779671
```

## data\$C11t	0.1979428568	0.0432196439	4.5799280	0.13685431
## data\$C12t	0.1759634678	0.0381362593	4.6140726	0.13587216
## data\$C13t	0.1493480011	0.0330260694	4.5221246	0.13854924
## data\$C14t	0.1265979430	0.0279661340	4.5268303	0.13840972
## data\$C15t	0.1025279665	0.0230305850	4.4518177	0.14066732
## data\$C16t	0.0794644224	0.0182894774	4.3448165	0.14401593
## data\$C17t	0.0579150053	0.0138082121	4.1942436	0.14900239
## data\$C18t	0.0378242717	0.0096486916	3.9201452	0.15900592
## data\$C19t	0.0192527623	0.0058782707	3.2752425	0.18865029
## data\$C20t	0.0025910125	0.0026455236	0.9793950	0.50662682
## data\$C21t	-0.0106776320	0.0014928362	-7.1525810	0.08843239
## data\$C22t	-0.0215614916	0.0035274078	-6.1125600	0.10323494
## data\$C23t	-0.0336396836	0.0055414244	-6.0705843	0.10393619
## data\$C24t	-0.0421335994	0.0071620164	-5.8829241	0.10719029
## data\$C25t	-0.0450744847	0.0083593220	-5.3921221	0.11673849
## data\$C26t	-0.0499790849	0.0091429226	-5.4664233	0.11518638
## data\$C27t	-0.0507242820	0.0095373650	-5.3184797	0.11831819
## data\$C28t	-0.0508480716	0.0095768130	-5.3094982	0.11851374
## data\$C29t	-0.0482090244	0.0093027237	-5.1822484	0.12135466
## data\$C30t	-0.0454170980	0.0087621870	-5.1833062	0.12133049
## data\$C31t	-0.0386270841	0.0080064170	-4.8245156	0.13011271
## data\$C32t	-0.0365837532	0.0070892650	-5.1604438	0.12185502
## data\$C33t	-0.0343000137	0.0060657359	-5.6547160	0.11143002
## data\$C34t	-0.0270168737	0.0049905228	-5.4136359	0.11628484
## data\$C35t	-0.0222807231	0.0039165953	-5.6887990	0.11077585
## data\$C36t	-0.0136207997	0.0028939017	-4.7067250	0.13327571
## data\$C37t	-0.0082871107	0.0019683240	-4.2102371	0.14845670
## data\$C38t	-0.0113945300	0.0011813819	-9.6450858	0.06576958
## data\$C39t	-0.0043520750	0.0005736575	-7.5865384	0.08343341
## data\$S1t	-0.0276725748	0.0058034704	-4.7682805	0.13160410
## data\$S2t	-0.0555469190	0.0115208032	-4.8214450	0.13019329
## data\$S3t	-0.0825828834	0.0170744228	-4.8366427	0.12979545
## data\$S4t	-0.1076686555	0.0223856057	-4.8097272	0.13050167
## data\$S5t	-0.1305477833	0.0273807484	-4.7678676	0.13161518
## data\$S6t	-0.1535240367	0.0319924273	-4.7987618	0.13079156
## data\$S7t	-0.1703570183	0.0361606221	-4.7111197	0.13315498
## data\$S8t	-0.1900763358	0.0398338159	-4.7717330	0.13151157
## data\$S9t	-0.2063351301	0.0429699190	-4.8018506	0.13070978
## data\$S10t	-0.2153434256	0.0455369893	-4.7289781	0.13266659
## data\$S11t	-0.2274505052	0.0475137355	-4.7870474	0.13110266
## data\$S12t	-0.2331468760	0.0488897887	-4.7688256	0.13158949
## data\$S13t	-0.2370475384	0.0496657409	-4.7728582	0.13148144
## data\$S14t	-0.2377232987	0.0498529474	-4.7684904	0.13159848
## data\$S15t	-0.2350219038	0.0494731009	-4.7504987	0.13208273
## data\$S16t	-0.2307953377	0.0485575869	-4.7530232	0.13201457
## data\$S17t	-0.2237293140	0.0471466360	-4.7453930	0.13222079
## data\$S18t	-0.2150462100	0.0452882937	-4.7483840	0.13213988
## data\$S19t	-0.2017394737	0.0430372317	-4.6875569	0.13380481
## data\$S20t	-0.1894674822	0.0404534284	-4.6835952	0.13391469
## data\$S21t	-0.1750467177	0.0376007492	-4.6554051	0.13470167
## data\$S22t	-0.1607127627	0.0345454593	-4.6522109	0.13479142
## data\$S23t	-0.1393484039	0.0313547042	-4.4442583	0.14089886
## data\$S24t	-0.1269441950	0.0280949910	-4.5183924	0.13866010
## data\$S25t	-0.1140094283	0.0248307073	-4.5914692	0.13652077

```
## data$S26t -0.0990847781 0.0216227103 -4.5824403 0.13678157
## data$S27t -0.0840724745 0.0185270193 -4.5378306 0.13808463
## data$S28t -0.0681187467 0.0155936408 -4.3683670 0.14326553
## data$S29t -0.0544291965 0.0128655553 -4.2306139 0.14776713
## data$S30t -0.0431806779 0.0103778882 -4.1608347 0.15015507
## data$S31t -0.0340256269 0.0081572887 -4.1711932 0.14979581
## data$S32t -0.0263468490 0.0062215342 -4.2347833 0.14762680
## data$S33t -0.0176601566 0.0045793813 -3.8564503 0.16152172
## data$S34t -0.0113030052 0.0032306864 -3.4986389 0.17723649
## data$S35t -0.0055395234 0.0021668227 -2.5565189 0.23736995
## data$S36t 0.0008266646 0.0013713918 0.6027925 0.65465271
## data$S37t 0.0002858443 0.0008208757 0.3482188 0.78667693
## data$S38t 0.0006705623 0.0004828485 1.3887632 0.39729273
## data$S39t 0.0007427796 0.0003055406 2.4310343 0.24844107
```

Notamos que ningún coeficiente es significativo usando éste modelo.

Ahora, analicemos la ANOVA del modelo forward:

```
modeloFordward$anova
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## data$y_training ~ 1
##
## Final Model:
## data$y_training ~ data$S23t + data$C38t + data$S36t + data$C33t +
## data$C10t + data$S24t + data$C31t + data$S7t + data$C35t +
## data$S10t + data$C12t + data$S29t + data$C39t + data$C24t +
## data$C34t + data$C22t + data$C14t + data$S32t + data$C3t +
## data$C37t + data$S30t + data$S28t + data$S19t + data$S21t +
## data$C36t + data$C15t + data$S37t + data$S35t + data$S18t +
## data$S9t + data$C25t + data$C1t + data$S20t + data$S3t +
## data$C16t + data$C17t + data$S22t + data$C21t + data$S5t +
## data$C18t + data$C29t + data$S8t + data$S26t + data$C8t +
## data$C2t + data$S27t + data$C19t + data$C30t + data$C27t +
## data$S15t + data$S12t + data$C4t + data$C11t + data$C5t +
## data$S33t + data$S39t + data$C26t + data$C23t + data$C7t
##
##
##          Step Df      Deviance Resid. Df  Resid. Dev      AIC
## 1
## 2 + data$S23t  1 1.619658e-03      78 6.060642e-03 -755.0373
## 3 + data$C38t  1 1.207336e-03      77 4.853306e-03 -770.8097
## 4 + data$S36t  1 6.075719e-04      76 4.245734e-03 -779.5094
## 5 + data$C33t  1 4.439680e-04      75 3.801766e-03 -786.3453
## 6 + data$C10t  1 3.838754e-04      74 3.417890e-03 -792.8607
## 7 + data$S24t  1 3.296221e-04      73 3.088268e-03 -798.9737
## 8 + data$C31t  1 2.911651e-04      72 2.797103e-03 -804.8958
## 9 + data$S7t   1 2.770916e-04      71 2.520012e-03 -811.2415
## 10 + data$C35t 1 2.322689e-04      70 2.287743e-03 -816.9773
## 11 + data$S10t 1 1.884928e-04      69 2.099250e-03 -821.8561
## 12 + data$C12t 1 1.697296e-04      68 1.929520e-03 -826.6008
## 13 + data$S29t 1 1.630208e-04      67 1.766500e-03 -831.6626
```


## 14 + data\$C39t	1	1.282125e-04	66	1.638287e-03	-835.6905
## 15 + data\$C24t	1	1.258545e-04	65	1.512432e-03	-840.0850
## 16 + data\$C34t	1	1.284207e-04	64	1.384012e-03	-845.1836
## 17 + data\$C22t	1	9.357495e-05	63	1.290437e-03	-848.7841
## 18 + data\$C14t	1	9.484221e-05	62	1.195595e-03	-852.8911
## 19 + data\$S32t	1	9.889671e-05	61	1.096698e-03	-857.7983
## 20 + data\$C3t	1	7.945347e-05	60	1.017244e-03	-861.8148
## 21 + data\$C37t	1	6.767255e-05	59	9.495719e-04	-865.3221
## 22 + data\$S30t	1	6.505970e-05	58	8.845122e-04	-869.0001
## 23 + data\$S28t	1	6.810052e-05	57	8.164117e-04	-873.4095
## 24 + data\$S19t	1	5.288636e-05	56	7.635253e-04	-876.7673
## 25 + data\$S21t	1	5.566720e-05	55	7.078581e-04	-880.8235
## 26 + data\$C36t	1	5.065325e-05	54	6.572048e-04	-884.7633
## 27 + data\$C15t	1	4.863945e-05	53	6.085654e-04	-888.9146
## 28 + data\$S37t	1	4.513706e-05	52	5.634283e-04	-893.0798
## 29 + data\$S35t	1	4.290608e-05	51	5.205222e-04	-897.4164
## 30 + data\$S18t	1	3.759597e-05	50	4.829263e-04	-901.4139
## 31 + data\$S9t	1	3.665751e-05	49	4.462688e-04	-905.7293
## 32 + data\$C25t	1	2.957723e-05	48	4.166915e-04	-909.2153
## 33 + data\$C1t	1	2.929477e-05	47	3.873968e-04	-913.0470
## 34 + data\$S20t	1	2.860577e-05	46	3.587910e-04	-917.1838
## 35 + data\$S3t	1	2.345314e-05	45	3.353379e-04	-920.5919
## 36 + data\$C16t	1	2.272486e-05	44	3.126130e-04	-924.2057
## 37 + data\$C17t	1	2.076172e-05	43	2.918513e-04	-927.7034
## 38 + data\$S22t	1	1.939977e-05	42	2.724515e-04	-931.2061
## 39 + data\$C21t	1	1.862843e-05	41	2.538231e-04	-934.8720
## 40 + data\$S5t	1	1.847773e-05	40	2.353454e-04	-938.9187
## 41 + data\$C18t	1	1.877779e-05	39	2.165676e-04	-943.5708
## 42 + data\$C29t	1	1.948650e-05	38	1.970811e-04	-949.1138
## 43 + data\$S8t	1	1.724822e-05	37	1.798329e-04	-954.4407
## 44 + data\$S26t	1	9.558292e-06	36	1.702746e-04	-956.8100
## 45 + data\$C8t	1	8.853546e-06	35	1.614210e-04	-959.0817
## 46 + data\$C2t	1	8.382294e-06	34	1.530387e-04	-961.3477
## 47 + data\$S27t	1	7.927517e-06	33	1.451112e-04	-963.6029
## 48 + data\$C19t	1	7.955575e-06	32	1.371556e-04	-966.1137
## 49 + data\$C30t	1	7.338544e-06	31	1.298171e-04	-968.5129
## 50 + data\$C27t	1	6.528065e-06	30	1.232890e-04	-970.6405
## 51 + data\$S15t	1	6.531270e-06	29	1.167578e-04	-972.9949
## 52 + data\$S12t	1	6.862003e-06	28	1.098958e-04	-975.8404
## 53 + data\$C4t	1	6.298912e-06	27	1.035968e-04	-978.5624
## 54 + data\$C11t	1	6.375774e-06	26	9.722106e-05	-981.6440
## 55 + data\$C5t	1	6.727549e-06	25	9.049352e-05	-985.3807
## 56 + data\$S33t	1	5.196474e-06	24	8.529704e-05	-988.1118
## 57 + data\$S39t	1	4.173232e-06	23	8.112381e-05	-990.1249
## 58 + data\$C26t	1	2.694154e-06	22	7.842966e-05	-990.8268
## 59 + data\$C23t	1	2.348872e-06	21	7.608078e-05	-991.2593
## 60 + data\$C7t	1	2.700314e-06	20	7.338047e-05	-992.1503

En la ANOVA estamos observando las covariables con las que sí se queda nuestro modelo, por lo que podemos notar que no está utilizando 19.

Analizaremos los coeficientes con los que se quedó al final el modelo:

```
summary(modeloFordward)
```

```
##
```

```
## Call:
## lm(formula = data$y_training ~ data$S23t + data$C38t + data$S36t +
##     data$C33t + data$C10t + data$S24t + data$C31t + data$S7t +
##     data$C35t + data$S10t + data$C12t + data$S29t + data$C39t +
##     data$C24t + data$C34t + data$C22t + data$C14t + data$S32t +
##     data$C3t + data$C37t + data$S30t + data$S28t + data$S19t +
##     data$S21t + data$C36t + data$C15t + data$S37t + data$S35t +
##     data$S18t + data$S9t + data$C25t + data$C1t + data$S20t +
##     data$S3t + data$C16t + data$C17t + data$S22t + data$C21t +
##     data$S5t + data$C18t + data$C29t + data$S8t + data$S26t +
##     data$C8t + data$C2t + data$S27t + data$C19t + data$C30t +
##     data$C27t + data$S15t + data$S12t + data$C4t + data$C11t +
##     data$C5t + data$S33t + data$S39t + data$C26t + data$C23t +
##     data$C7t, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.340e-03 -6.032e-04  5.040e-06  6.756e-04  1.831e-03
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0004959  0.0002258  -2.197  0.040003 *
## data$S23t    0.0065426  0.0003133  20.883 4.73e-15 ***
## data$C38t   -0.0054002  0.0003330 -16.218 5.65e-13 ***
## data$S36t    0.0037394  0.0003080  12.141 1.11e-10 ***
## data$C33t   -0.0029545  0.0003221  -9.174 1.32e-08 ***
## data$C10t   -0.0033191  0.0003253 -10.203 2.25e-09 ***
## data$S24t    0.0030561  0.0003147   9.713 5.15e-09 ***
## data$C31t    0.0030346  0.0003200   9.482 7.68e-09 ***
## data$S7t     0.0030985  0.0003107   9.972 3.31e-09 ***
## data$C35t   -0.0021387  0.0003261  -6.559 2.16e-06 ***
## data$S10t    0.0026788  0.0003164   8.467 4.80e-08 ***
## data$C12t    0.0019902  0.0003252   6.120 5.57e-06 ***
## data$S29t    0.0019922  0.0003205   6.215 4.53e-06 ***
## data$C39t   -0.0015612  0.0003348  -4.664 0.000149 ***
## data$C24t   -0.0012261  0.0003290  -3.726 0.001333 **
## data$C34t   -0.0012942  0.0003239  -3.996 0.000710 ***
## data$C22t    0.0020277  0.0003319   6.109 5.71e-06 ***
## data$C14t    0.0015941  0.0003269   4.876 9.15e-05 ***
## data$S32t   -0.0014047  0.0003184  -4.412 0.000269 ***
## data$C3t    -0.0017862  0.0003372  -5.297 3.49e-05 ***
## data$C37t    0.0017810  0.0003308   5.384 2.86e-05 ***
## data$S30t    0.0013776  0.0003205   4.299 0.000350 ***
## data$S28t    0.0013804  0.0003200   4.313 0.000338 ***
## data$S19t    0.0015040  0.0003117   4.826 0.000103 ***
## data$S21t    0.0014174  0.0003116   4.549 0.000195 ***
## data$C36t    0.0012284  0.0003284   3.740 0.001290 **
## data$C15t    0.0012522  0.0003283   3.815 0.001085 **
## data$S37t    0.0009259  0.0003050   3.036 0.006527 **
## data$S35t    0.0008283  0.0003111   2.663 0.014938 *
## data$S18t   -0.0006390  0.0003123  -2.046 0.054146 .
## data$S9t    -0.0004546  0.0003150  -1.443 0.164387
## data$C25t    0.0012288  0.0003272   3.756 0.001243 **
## data$C1t    -0.0013451  0.0003395  -3.962 0.000768 ***
```

```
## data$S20t    0.0010237  0.0003114   3.288 0.003680 **
## data$S3t    -0.0005726  0.0003006  -1.905 0.071307 .
## data$C16t   0.0009583  0.0003297   2.907 0.008717 **
## data$C17t   0.0009147  0.0003310   2.764 0.011978 *
## data$S22t   0.0007748  0.0003122   2.481 0.022090 *
## data$C21t   0.0009248  0.0003327   2.780 0.011568 *
## data$S5t    0.0008992  0.0003054   2.944 0.008021 **
## data$C18t   0.0007957  0.0003320   2.396 0.026455 *
## data$C29t   0.0007205  0.0003206   2.248 0.036047 *
## data$S8t    0.0008972  0.0003130   2.866 0.009548 **
## data$S26t  -0.0005047  0.0003177  -1.589 0.127790
## data$C8t    -0.0007346  0.0003276  -2.243 0.036404 *
## data$C2t    -0.0007902  0.0003386  -2.334 0.030166 *
## data$S27t  -0.0004509  0.0003190  -1.413 0.172934
## data$C19t   0.0004315  0.0003328   1.297 0.209482
## data$C30t   0.0003941  0.0003200   1.232 0.232401
## data$C27t   0.0003450  0.0003233   1.067 0.298691
## data$S15t   0.0005298  0.0003156   1.679 0.108773
## data$S12t   0.0005090  0.0003175   1.603 0.124611
## data$C4t    -0.0006200  0.0003355  -1.848 0.079421 .
## data$C11t  -0.0005403  0.0003250  -1.663 0.111960
## data$C5t    -0.0005172  0.0003334  -1.551 0.136553
## data$S33t  -0.0003493  0.0003164  -1.104 0.282642
## data$S39t  -0.0003053  0.0002999  -1.018 0.320950
## data$C26t  -0.0003360  0.0003252  -1.033 0.313866
## data$C23t  -0.0002849  0.0003307  -0.862 0.399077
## data$C7t    -0.0002825  0.0003293  -0.858 0.401117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.001915 on 20 degrees of freedom
## Multiple R-squared:  0.9904, Adjusted R-squared:  0.9623
## F-statistic: 35.14 on 59 and 20 DF,  p-value: 2.714e-12
```

Notemos que claramente el modelo hecho por el método backward es el que tiene menor AIC, sin embargo, eso no significa que sea el mejor modelo, ya que, tiene menor AIC que el modelo hecho con el método forward, porque se está tomando muchas más covariables, de hecho toma todas, mientras que por otro lado el del forward, toma muchas menos covariables y sí, tiene menor AIC, pero observando su R^2 no tiene prácticamente error, por lo que seleccionaré este último como mejor modelo.

Analizando los coeficientes del modelo forward, nos damos cuenta que tiene variables que no son significativas para mi ajuste, por lo que no tomaremos en cuenta las covariables que su p-value supere mi nivel de significancia. Es importante recalcar que al quitar una covariables puede cambiar el valor del p-value de las demás, por lo que vamos a tener que aplicar ahora un modelo backward, el cual, consistirá en partiendo de las covariables con las que nos ha arrojado el modelo forward, irá quitando la covariable menos significativa (la del mayor p-value) hasta quedaros con puras variables que su p-value $< 5\%$.

Éste proceso lo haré en la consola, por lo que no se mostrarán los resultados, lo hago así para que el pdf no se extienda demasiado.

Al final el modelo me queda:

```
# Actualizo a mi modelo final
modeloFordward <- update(modeloFordward, data$y_training ~ data$S23t + data$C38t + data$S36t + data$C33t
summary(modeloFordward)
```

```
##
## Call:
## lm(formula = data$y_training ~ data$S23t + data$C38t + data$S36t +
##     data$C33t + data$C10t + data$S24t + data$C31t + data$S7t +
##     data$C35t + data$S10t + data$C12t + data$S29t + data$C39t +
##     data$C24t + data$C34t + data$C22t + data$C14t + data$S32t +
##     data$C3t + data$C37t + data$S30t + data$S28t + data$S19t +
##     data$S21t + data$C36t + data$C15t + data$S37t + data$S35t,
##     data = data)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.0073866	-0.0010643	0.0001569	0.0020554	0.0058279

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.0002627	0.0003592	-0.731	0.467924
data\$S23t	0.0065019	0.0005093	12.767	< 2e-16 ***
data\$C38t	-0.0052913	0.0005326	-9.935	1.61e-13 ***
data\$S36t	0.0039300	0.0005047	7.786	3.14e-10 ***
data\$C33t	-0.0030200	0.0005181	-5.829	3.79e-07 ***
data\$C10t	-0.0032202	0.0005088	-6.329	6.23e-08 ***
data\$S24t	0.0030655	0.0005107	6.002	2.04e-07 ***
data\$C31t	0.0028820	0.0005133	5.614	8.19e-07 ***
data\$S7t	0.0027187	0.0005038	5.396	1.78e-06 ***
data\$C35t	-0.0021230	0.0005240	-4.051	0.000174 ***
data\$S10t	0.0022270	0.0005071	4.392	5.69e-05 ***
data\$C12t	0.0019681	0.0005083	3.872	0.000309 ***
data\$S29t	0.0021904	0.0005159	4.246	9.23e-05 ***
data\$C39t	-0.0014322	0.0005347	-2.679	0.009926 **
data\$C24t	-0.0016032	0.0005112	-3.136	0.002842 **
data\$C34t	-0.0013178	0.0005210	-2.529	0.014557 *
data\$C22t	0.0016395	0.0005124	3.200	0.002366 **
data\$C14t	0.0014564	0.0005092	2.860	0.006120 **
data\$S32t	-0.0011678	0.0005137	-2.273	0.027264 *
data\$C3t	-0.0013588	0.0005180	-2.623	0.011453 *
data\$C37t	0.0018638	0.0005300	3.516	0.000928 ***
data\$S30t	0.0015958	0.0005157	3.094	0.003200 **
data\$S28t	0.0015518	0.0005156	3.010	0.004055 **
data\$S19t	0.0012575	0.0005058	2.486	0.016235 *
data\$S21t	0.0012725	0.0005069	2.510	0.015282 *
data\$C36t	0.0012800	0.0005271	2.428	0.018739 *
data\$C15t	0.0010617	0.0005099	2.082	0.042369 *
data\$S37t	0.0010898	0.0005022	2.170	0.034692 *
data\$S35t	0.0010402	0.0005073	2.050	0.045488 *

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.003195 on 51 degrees of freedom
## Multiple R-squared:  0.9322, Adjusted R-squared:  0.895
## F-statistic: 25.05 on 28 and 51 DF, p-value: < 2.2e-16

# Error cuadrático
print(paste("La R cuadrada del mejor modelo es:", round(summary(modeloFordward)$r.squared, 5)))
```

```
## [1] "La R cuadrada del mejor modelo es: 0.93223"
```

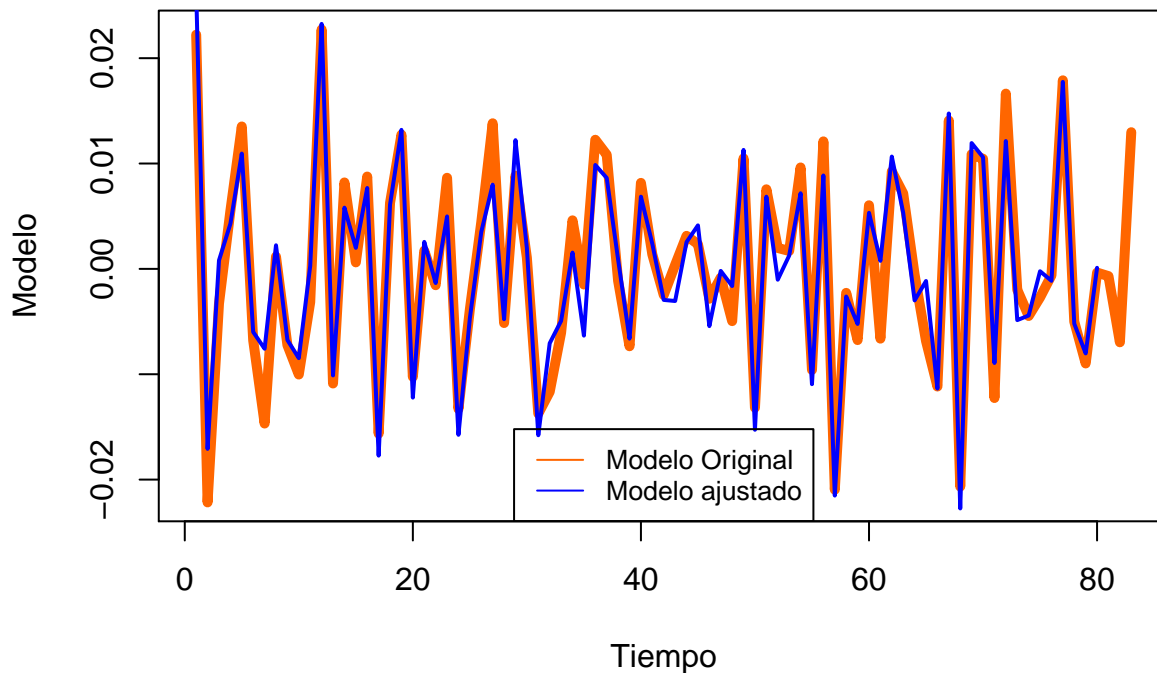
Como podemos observar, ya son significativas todas las covariables.

Presento la gráfica del ajuste.

```
# Generamos la gráfica
```

```
plot(DDLnZt, type = "l", main = "Modelo Original vs Modelo Ajustado", col="#FF6600",lwd=5, xlab = "Tiempo")
lines(modeloFordward$fitted.values,col="blue",lwd=2, lty = 1)
legend("bottom",legend = c("Modelo Original","Modelo ajustado"), col = c("#FF6600","blue"), lty=c(1,1),
```

Modelo Original vs Modelo Ajustado



Finalmente, vamos a realizar las **proyecciones** para las 3 últimas observaciones:

```
# Obtenemos los coeficientes del mejor modelo
```

```
coeficientes <- as.matrix(modeloFordward$coefficients)
```

```
# Extraemos los valores de los senos y los cosenos para las observaciones que vamos a predecir
```

```
covariables <- matriz[81:83, c("S23t", "C38t", "S36t", "C33t", "C10t", "S24t", "C31t", "S7t", "C35t", "
```

```
# Inicializo mi vector de predicciones
```

```
predict <- c()
```

```
# Lleno el vector
```

```
for (i in 1:3){
  predict[i+1]<-sum(t(coeficientes[-1])*as.matrix(covariables[i,]))+coeficientes[1] }
```

```
# Para que las graficas se junten, le agregamos un valor al vector de las predicciones
```

```
predict[1] <- modeloFordward$fitted.values[80]
```

```
# Graficamos el Ajuste, la predicción y la serie original
```

```
plot(DDLnZt, type = "l", main = "Modelo Original y valores Pronosticados", col="#FF6600",lwd=5, xlab = "Tiempo")
lines(modeloFordward$fitted.values,col="blue",lwd=2, lty = 2)
```

```
lines(y = predict, x = rep(80:83), col = "darkorchid4", lwd=3, lty = 1)
legend("bottom", legend = c("Modelo Original", "Modelo ajustado", "Pronósticos"), col = c("#FF6600", "blue", "darkorchid4"), lty = c(1, 2, 1))
```

Modelo Original y valores Pronosticados

