A8-Series de tiempo

Erika Martínez Meneses

2024-11-12

Para los datos de las ventas de televisores analiza la serie de tiempo más apropiada:

```
# Crear tabla
datos <- data.frame(
   Año = c(1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4),
   Trimestre = c(1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4),
   Ventas = c(4.8, 4.1, 6.0, 6.5, 5.8, 5.2, 6.8, 7.4, 6.0, 5.6, 7.5, 7.8,
6.3, 5.9, 8.0, 8.4)
)</pre>
```

Realiza el análisis de tendencia y estacionalidad

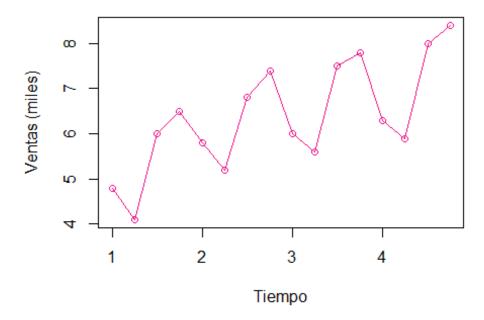
Identifica si es una serie estacionaria

```
# Convertir los datos a serie de tiempo
ventas_ts <- ts(datos$Ventas, start = c(1, 1), frequency = 4)</pre>
```

Grafica la serie para verificar su tendencia y estacionalidad

```
plot(ventas_ts, main = "Ventas trimestrales", xlab = "Tiempo", ylab = "Ve
ntas (miles)", col = "deeppink2", type = "o")
```

Ventas trimestrales



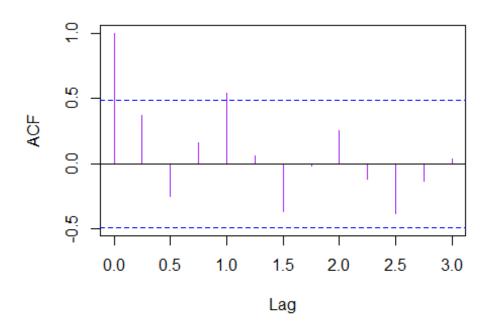
La gráfica muestra las ventas trimestrales de televisores en miles a lo largo de cuatro años. Se observa una tendencia ascendente en las ventas, con un patrón estacional en cada año. Las ventas tienden a subir y bajar de forma regular dentro de cada ciclo anual, con un comportamiento consistente en los trimestres. Este patrón sugiere la presencia de estacionalidad en los datos, ya que los aumentos y disminuciones de ventas se repiten de forma similar en cada año. Además, la tendencia general es de crecimiento, ya que los picos de cada año tienden a ser más altos que los del año anterior, lo cual indica un aumento en la demanda de televisores a lo largo del tiempo.

La gráfica también sugiere que la serie de tiempo podría no ser estacionaria, ya que tanto la media como la varianza parecen cambiar a lo largo del tiempo, con una tendencia creciente en las ventas y variaciones estacionales recurrentes. Esto indica que el comportamiento de las ventas no es constante y depende del período, por lo que a continuación se realizan pruebas adicionales para confirmar la no estacionariedad.

Analiza su gráfico de autocorrelación

acf(ventas_ts, col = "purple", main = "Gráfico de Autocorrelación")

Gráfico de Autocorrelación



qnorm(1-0.05/2)/sqrt(length(datos\$Ventas))
[1] 0.489991

El gráfico de autocorrelación (ACF) muestra los coeficientes de autocorrelación de las ventas trimestrales en diferentes rezagos (lags). Observamos una autocorrelación significativa en el primer rezago, que excede el valor crítico de \$\$0.489991 (calculado para un intervalo de confianza del 95%). Esto indica que las ventas de un trimestre están fuertemente correlacionadas con las del trimestre anterior, lo que sugiere una dependencia temporal en los datos. En rezagos mayores, las autocorrelaciones disminuyen y oscilan alrededor de cero, lo cual es típico en una serie con estacionalidad o tendencia. Este comportamiento refuerza la idea de que la serie podría no ser estacionaria y que existe una estructura temporal que podría modelarse adecuadamente con técnicas de series de tiempo, como un modelo ARIMA, para capturar la tendencia y estacionalidad. En el cuarto rezago vemos que excede nuevamente el valor crítico por lo que realizaremos un análisis de estos dos lags.

Análisis de estacionariedad

Prueba de hipótesis (Prueba de Dickey-Fuller Aumentada)

 H_0 : La serie no es estacionaria. ($\phi = 1$) H_1 : La serie es estacionaria.

Regla de decisión: Se rechaza H_0 si valor p < $\alpha = 0.05$

library(tseries)

```
## Warning: package 'tseries' was built under R version 4.4.2
## Registered S3 method overwritten by 'quantmod':
## method from
## as.zoo.data.frame zoo

adf_test <- adf.test(ventas_ts, k=4)
adf_test
##
## Augmented Dickey-Fuller Test
##
## data: ventas_ts
## data: ventas_ts
## Dickey-Fuller = -4.1942, Lag order = 4, p-value = 0.01648
## alternative hypothesis: stationary</pre>
```

con K=4 se rechaza H_0 indicando que la serie es estacionaria. Aunque este resultado sugiere estacionariedad, es posible que esta estacionariedad esté influenciada por la inclusión de los rezagos y no refleje completamente la tendencia y estacionalidad que observamos en la gráfica. La prueba podría estar capturando patrones a corto plazo que no eliminan completamente la tendencia subyacente de la serie.

```
library(forecast)
## Warning: package 'forecast' was built under R version 4.4.2
auto.arima(ventas ts)
## Series: ventas ts
## ARIMA(0,1,0)(0,1,0)[4]
##
## sigma^2 = 0.08001: log likelihood = -1.72
## AIC=5.43 AICc=5.88
                          BIC=5.83
library(tseries)
adf_test <- adf.test(ventas_ts, k=0)</pre>
adf_test
##
   Augmented Dickey-Fuller Test
##
##
## data: ventas ts
## Dickey-Fuller = -3.2388, Lag order = 0, p-value = 0.1004
## alternative hypothesis: stationary
```

Realizamos nuevamente la prueba con k=0 y se acepta H_0 , indicando que la serie no es estacionaria. Este resultado indica que sin considerar rezagos adicionales, no se puede rechazar la hipótesis de no estacionariedad, lo que coincide más con la observación visual de que la serie no es estacionaria en su forma original.

La recomendación de auto.arima para diferenciar la serie (ARIMA(0,1,0)(0,1,0)[4]) confirma la no estacionariedad inicial de la serie y sugiere que una primera diferencia

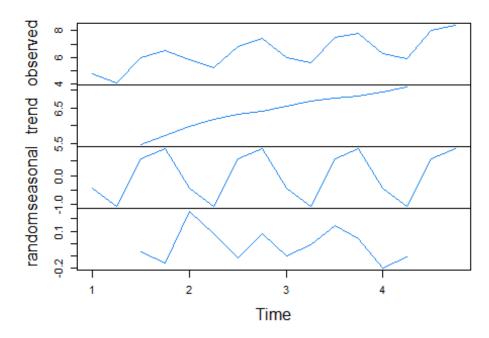
regular y estacional puede ayudar a estabilizar la tendencia y eliminar la estacionalidad. Esto también se alinea con la indicación visual de la serie de que necesita ser transformada para lograr estacionariedad.

En conclusión, aunque el resultado de la prueba de Dickey-Fuller con rezagos podría sugerir estacionariedad en el corto plazo, tanto la visualización como los otros análisis indican que la serie es no estacionaria en su forma original y necesita transformaciones para lograr estacionariedad.

Identifica si el modelo puede ser sumativo o multiplicativo (puedes probar con ambos para ver con cuál es mejor el modelo)

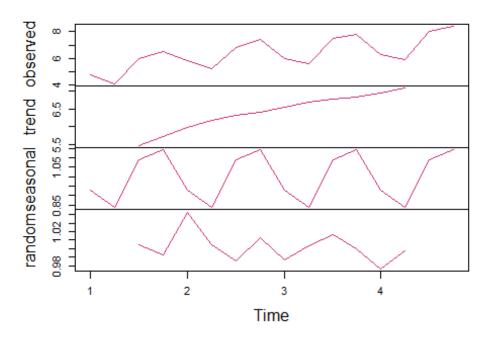
```
# Modelo multiplicativo vs. sumativo
decomp_add <- decompose(ventas_ts, type = "additive")
decomp_mult <- decompose(ventas_ts, type = "multiplicative")
plot(decomp_add, col = "#1C86EE")</pre>
```

Decomposition of additive time series



```
plot(decomp_mult, col = "#CD3278")
```

Decomposition of multiplicative time series



En la descomposición aditiva, la serie observada se divide en tres componentes: tendencia, estacionalidad y componente aleatorio (residual). La tendencia muestra un incremento constante a lo largo del tiempo, mientras que la estacionalidad presenta fluctuaciones regulares y uniformes en cada período. Este comportamiento es característico de un modelo aditivo, en el cual los componentes se suman para formar la serie observada. Mientras que en la descomposición multiplicativa, la estructura de la serie de tiempo se modela de forma que cada componente se multiplica para formar la serie observada. Aunque las tendencias y estacionalidades son similares a las del modelo aditivo, la estacionalidad y la variabilidad en los componentes aleatorios parecen depender del nivel de la tendencia, lo cual es característico de un modelo multiplicativo.

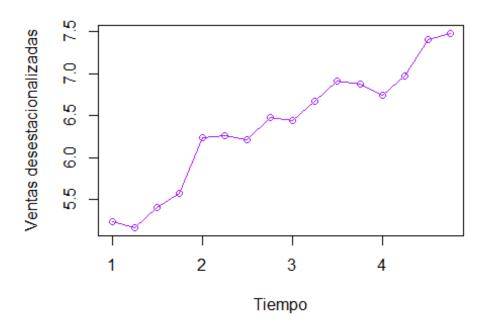
Calcula los índices estacionales y grafica la serie desestacionalizada

Se realizará para ambos modelos, multiplicatico y aditivo

Aditivo

```
ventas_desestacionalizadas <- ventas_ts - decomp_add$seasonal
plot(ventas_desestacionalizadas, main = "Ventas Desestacionalizadas", xla
b = "Tiempo", ylab = "Ventas desestacionalizadas", col = "purple", type =
"o")</pre>
```

Ventas Desestacionalizadas



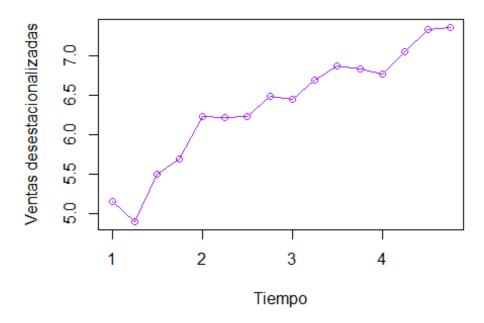
La gráfica de las ventas desestacionalizadas muestra una tendencia creciente en las ventas a lo largo del tiempo, eliminando la variabilidad estacional y permitiendo observar la tendencia con mayor claridad. Esto confirma que, independientemente de los cambios estacionales, las ventas de televisores han estado aumentando de manera consistente durante el período analizado.

En cuanto a los índices estacionales, estos indican el efecto promedio de la estacionalidad en cada trimestre. Estos valores sugieren que las ventas son más bajas en los primeros dos trimestres del año, especialmente en el segundo trimestre, y que tienden a ser mayores en los últimos dos trimestres, alcanzando su punto más alto en el cuarto trimestre. Este comportamiento estacional podría estar influenciado por factores como eventos de fin de año o cambios en la demanda estacional de televisores.

Multiplicativo

```
ventas_desestacionalizadas2 <- ventas_ts / decomp_mult$seasonal
plot(ventas_desestacionalizadas2, main = "Ventas Desestacionalizadas (Mul
tiplicativo)", xlab = "Tiempo", ylab = "Ventas desestacionalizadas", col
= "purple", type = "o")</pre>
```

Ventas Desestacionalizadas (Multiplicativo)



La gráfica muestra las ventas desestacionalizadas utilizando un modelo multiplicativo, eliminando la influencia estacional para resaltar la tendencia subyacente en los datos. Se observa una tendencia ascendente en las ventas desestacionalizadas a lo largo del tiempo, lo cual indica un crecimiento constante en las ventas trimestre a trimestre. Este comportamiento sugiere que, excluyendo los efectos estacionales, la demanda del producto o servicio representado sigue aumentando.

Analiza el modelo lineal de la tendencia

Aditivo

```
# Crear variable de tiempo para la regresión
tiempo <- 1:length(ventas desestacionalizadas)</pre>
# Ajuste del modelo lineal
modelo_lineal <- lm(ventas_desestacionalizadas ~ tiempo)</pre>
summary(modelo_lineal)
##
## Call:
## lm(formula = ventas_desestacionalizadas ~ tiempo)
##
## Residuals:
       Min
                1Q Median
                                 3Q
                                         Max
## -0.2992 -0.1486 -0.0037 0.1005 0.3698
##
```

```
## Coefficients:

## Estimate Std. Error t value Pr(>|t|)

## (Intercept) 5.13917 0.10172 50.52 < 2e-16 ***

## tiempo 0.14613 0.01052 13.89 1.4e-09 ***

## ---

## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##

## Residual standard error: 0.194 on 14 degrees of freedom

## Multiple R-squared: 0.9324, Adjusted R-squared: 0.9275

## F-statistic: 193 on 1 and 14 DF, p-value: 1.399e-09
```

El modelo lineal ajustado para la tendencia de las ventas desestacionalizadas tiene una ecuación de la forma: Ventas desestacionalizadas $= 5.13917 + 0.14613 \cdot tiempo$

El coeficiente de intercepto (5.13917) representa el valor estimado de las ventas desestacionalizadas al inicio del período analizado. El coeficiente de la variable tiempo (0.14613) es positivo y significativo (valor p = 1.4e-09), indicando que, en promedio, las ventas desestacionalizadas aumentan aproximadamente 0.146 unidades por cada periodo de tiempo. Este crecimiento constante sugiere una tendencia lineal creciente en las ventas de televisores.

El R-cuadrado múltiple de 0.9324 implica que el modelo lineal explica aproximadamente el 93.24% de la variabilidad en las ventas desestacionalizadas, lo que indica un buen ajuste. Además, el error estándar de los residuos es bajo (0.194), lo que sugiere que los valores predichos están relativamente cerca de los valores observados.

Multiplicativo

```
tiempo2 <- 1:length(ventas desestacionalizadas2)</pre>
modelo_lineal2 <- lm(ventas_desestacionalizadas2 ~ tiempo2)</pre>
summary(modelo_lineal2)
##
## Call:
## lm(formula = ventas desestacionalizadas2 ~ tiempo2)
## Residuals:
##
       Min
                10 Median
                                30
                                       Max
## -0.5007 -0.1001 0.0037 0.1207 0.3872
##
## Coefficients:
##
              Estimate Std. Error t value Pr(>|t|)
                                     45.73 < 2e-16 ***
## (Intercept) 5.10804
                           0.11171
                                     12.76 4.25e-09 ***
## tiempo2
                0.14738
                           0.01155
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.213 on 14 degrees of freedom
```

```
## Multiple R-squared: 0.9208, Adjusted R-squared: 0.9151
## F-statistic: 162.7 on 1 and 14 DF, p-value: 4.248e-09
```

Ecuación de la tendencia: $Ventas = 5.108 + 0.147 \cdot Tiempo$

Al comparar el modelo de tendencia lineal aditivo y el multiplicativo, encontramos que ambos ofrecen un ajuste fuerte a los datos, con valores de R-cuadrado ajustado cercanos a 1, lo que indica que explican bien la variabilidad de las ventas trimestrales. Sin embargo, el modelo aditivo muestra un ajuste ligeramente superior, con un R-cuadrado ajustado de 0.9275 en comparación con el 0.9151 del modelo multiplicativo. Además, el error estándar de los residuos es menor en el modelo aditivo (0.194 frente a 0.213), lo que sugiere una mayor precisión en las predicciones dentro de la muestra.

Ambos modelos presentan una alta significancia estadística en sus coeficientes, con valores p extremadamente bajos para la variable de tiempo, indicando que la tendencia capturada por ambos modelos es significativa. Sin embargo, el modelo aditivo también tiene un valor p marginalmente mejor.

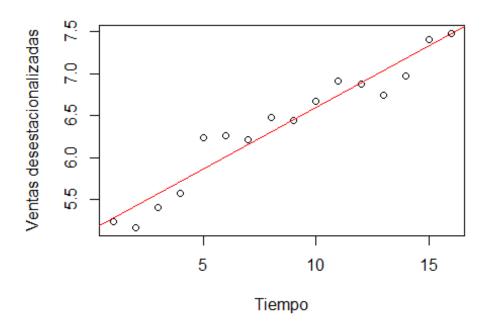
En conclusión, dado el mejor ajuste y precisión del modelo aditivo, parece ser la opción preferible para describir la tendencia en las ventas trimestrales en este conjunto de datos. Este modelo proporciona una mejor explicación de la variabilidad observada y genera predicciones más precisas, lo que lo hace más adecuado para representar la serie temporal de ventas.

Realiza la regresión lineal de la tendencia (ventas desestacionalizadas vs tiempo)

Se continua con el modelo aditivo.

```
plot(tiempo, ventas_desestacionalizadas, main = "Modelo de tendencia line
al", xlab = "Tiempo", ylab = "Ventas desestacionalizadas")
abline(modelo lineal, col = "red")
```

Modelo de tendencia lineal



La gráfica muestra el modelo de tendencia lineal ajustado para las ventas desestacionalizadas de televisores en función del tiempo. Los puntos representan las ventas desestacionalizadas observadas en cada periodo, mientras que la línea roja representa la línea de regresión lineal ajustada.

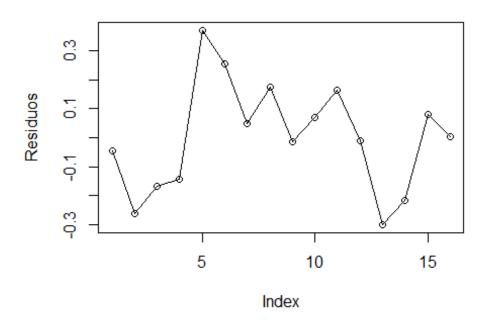
Se observa una tendencia creciente en las ventas desestacionalizadas, con la línea de regresión capturando esta tendencia de manera consistente. Esto confirma que, en promedio, las ventas de televisores han estado aumentando a lo largo del tiempo de forma lineal. La cercanía de los puntos a la línea de tendencia indica que el modelo lineal es adecuado para describir esta tendencia creciente en las ventas.

Analiza la significancia del modelo lineal, global e individual

Haz el análisis de residuos

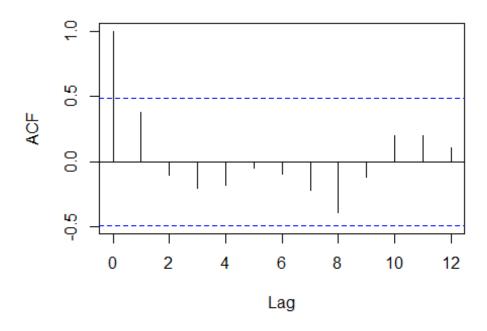
```
plot(modelo_lineal$residuals, main = "Residuos del Modelo Lineal", ylab =
"Residuos", type = "o")
```

Residuos del Modelo Lineal



acf(modelo_lineal\$residuals, main = "ACF de los Residuos del Modelo Linea
1")

ACF de los Residuos del Modelo Lineal



La primera gráfica muestra los residuos del modelo lineal a lo largo del tiempo. Los residuos representan la diferencia entre los valores observados de las ventas desestacionalizadas y los valores predichos por el modelo lineal. Se observa que los residuos están distribuidos en torno a cero, pero existe cierta variabilidad en su amplitud. La ausencia de un patrón claro en los residuos sugiere que el modelo lineal ha capturado la tendencia principal, aunque aún podría haber variaciones que el modelo no explica completamente.

La segunda gráfica muestra el correlograma de los residuos (ACF), que indica la autocorrelación de los residuos a diferentes rezagos. La mayoría de los valores de autocorrelación caen dentro de los límites de significancia (líneas punteadas azules), lo cual sugiere que los residuos son en su mayoría independientes y no presentan autocorrelación significativa. Esto es una buena señal, ya que sugiere que el modelo lineal ha capturado adecuadamente la estructura de la serie y que los residuos se comportan de manera aleatoria, como se esperaría en un modelo bien ajustado. Sin embargo, un ligero valor de autocorrelación en el primer rezago podría indicar la presencia de alguna pequeña dependencia temporal no capturada.

Calcula el CME y el EPAM de la predicción de la serie de tiempo

```
predicciones_lineales <- predict(modelo_lineal)
CME <- mean((ventas_desestacionalizadas - predicciones_lineales)^2)
EPAM <- mean(abs((ventas_desestacionalizadas-predicciones_lineales)/venta
s_desestacionalizadas))*100
CME
## [1] 0.03291917
EPAM
## [1] 2.341319</pre>
```

Los resultados de las métricas de error para el modelo lineal ajustado indican que el Error Cuadrático Medio (CME) es de aproximadamente 0.0329. Este valor relativamente bajo sugiere que, en promedio, las desviaciones cuadráticas de las predicciones respecto a los valores observados son pequeñas, lo cual indica que el modelo tiene un buen nivel de precisión en sus estimaciones.

Por otro lado, el Error Porcentual Absoluto Medio (EPAM) es de 2.34%, lo cual significa que, en promedio, las predicciones del modelo difieren de los valores observados en un 2.34%. Este valor bajo también sugiere que el modelo lineal proporciona predicciones precisas y confiables para las ventas desestacionalizadas, con un margen de error relativamente pequeño en términos porcentuales.

Explora un mejor modelo, por ejemplo un modelo cuadrático: $y = \beta_0 + \beta_1 x + \beta_2 x^2$

Para ello transforma la variable ventas (recuerda que la regresión no lineal es una regresión lineal con una tranformación).

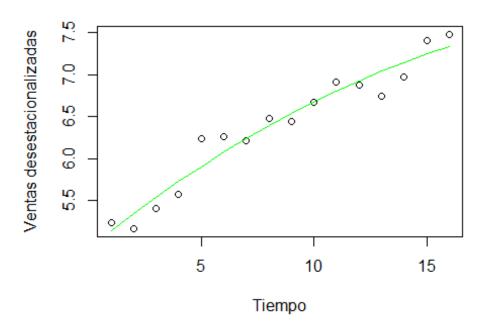
```
modelo_cuadratico <- lm(ventas_desestacionalizadas ~ tiempo + I(tiempo^2)</pre>
)
summary(modelo cuadratico)
##
## Call:
## lm(formula = ventas_desestacionalizadas ~ tiempo + I(tiempo^2))
## Residuals:
                      Median
##
       Min
                  10
                                    3Q
                                            Max
## -0.30333 -0.13440 -0.01928 0.11368 0.33301
##
## Coefficients:
                Estimate Std. Error t value Pr(>|t|)
                          0.155679 31.673 1.08e-13 ***
## (Intercept) 4.930833
## tiempo
                0.215572
                          0.042149 5.115 0.000199 ***
## I(tiempo^2) -0.004085  0.002410 -1.695 0.113918
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1822 on 13 degrees of freedom
## Multiple R-squared: 0.9446, Adjusted R-squared: 0.9361
## F-statistic: 110.8 on 2 and 13 DF, p-value: 6.805e-09
```

La regresión cuadrática sobre las ventas desestacionalizadas muestra que el modelo tiene un intercepto estimado de 4.9308, un coeficiente lineal de 0.2156 y un coeficiente cuadrático de -0.0041. Estos coeficientes sugieren una relación no estrictamente lineal entre el tiempo y las ventas desestacionalizadas, aunque la influencia cuadrática es leve y no es estadísticamente significativa al nivel del 5% (p = 0.114). Esto indica que el término cuadrático podría no mejorar sustancialmente el modelo en comparación con una tendencia lineal simple.

El R-cuadrado ajustado es de 0.9361, ligeramente superior al modelo lineal, lo que sugiere que el modelo cuadrático explica marginalmente más variabilidad en las ventas desestacionalizadas. Sin embargo, dado que el término cuadrático no es estadísticamente significativo y que la mejora en R-cuadrado es mínima, es posible que un modelo lineal simple ya sea adecuado.

```
# Graficar el modelo cuadrático junto con los datos desestacionalizados
plot(tiempo, ventas_desestacionalizadas, main = "Modelo de tendencia cuad
rática", xlab = "Tiempo", ylab = "Ventas desestacionalizadas")
lines(tiempo, predict(modelo_cuadratico), col = "green")
```

Modelo de tendencia cuadrática



En la gráfica, la línea verde representa el ajuste del modelo cuadrático, mostrando una ligera curvatura, lo cual podría capturar pequeñas desviaciones en la tendencia lineal, aunque no de forma considerable.

Concluye sobre el mejor modelo

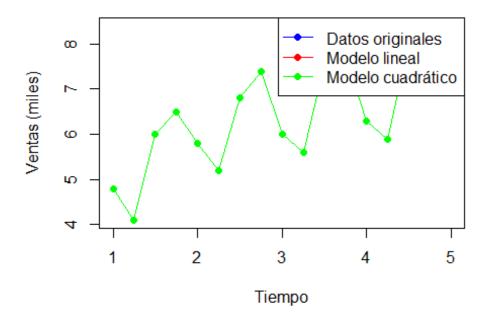
En conclusión, al analizar los modelos de tendencia lineal y cuadrática para las ventas desestacionalizadas, el modelo lineal parece ser la mejor opción. Aunque el modelo cuadrático muestra un ajuste ligeramente superior en términos de R-cuadrado ajustado, la contribución del término cuadrático no es estadísticamente significativa (p > 0.05). Esto sugiere que la mejora en el ajuste del modelo cuadrático no justifica la complejidad adicional que aporta el término cuadrático.

El modelo lineal es más sencillo, tiene un R-cuadrado ajustado de 0.9275, lo que indica que explica bien la variabilidad en las ventas desestacionalizadas, y presenta un error cuadrático medio (CME) y un error porcentual absoluto medio (EPAM) bajos. Por lo tanto, se recomienda utilizar el modelo de tendencia lineal para describir y proyectar la tendencia en las ventas desestacionalizadas, dado que proporciona un ajuste adecuado sin añadir términos innecesarios.

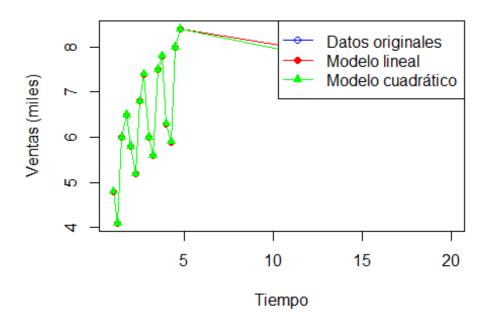
Realiza el pronóstico para el siguiente año y grafícalo junto con los pronósticos previos y los datos originales.

```
# Pronóstico para el siguiente año
nuevo_tiempo <- length(ventas_ts) + (1:4) # para los próximos 4 trimestre
s
prediccion_lineal <- predict(modelo_lineal, newdata = data.frame(tiempo = nuevo_tiempo))
prediccion_cuadratica <- predict(modelo_cuadratico, newdata = data.frame(tiempo = nuevo_tiempo))
# Graficar los pronósticos junto con los datos originales
plot(ventas_ts, main = "Pronóstico para el siguiente año", xlab = "Tiempo", ylab = "Ventas (miles)", col = "blue", type = "o", xlim = c(1, 5))
lines(c(time(ventas_ts), nuevo_tiempo), c(ventas_ts, prediccion_lineal),
col = "red", type = "o", pch = 16)
lines(c(time(ventas_ts), nuevo_tiempo), c(ventas_ts, prediccion_cuadratic
a), col = "green", type = "o", pch = 16)
legend("topright", legend = c("Datos originales", "Modelo lineal", "Model
o cuadrático"), col = c("blue", "red", "green"), lty = 1, pch = 16)</pre>
```

Pronóstico para el siguiente año



Pronóstico para el siguiente año



Observando las gráficas, ambos modelos (lineal en rojo y cuadrático en verde) están muy próximos entre sí y siguen de cerca la tendencia de los datos históricos, lo cual indica que ambos modelos ajustan los datos de manera similar en el corto plazo.

En la primera gráfica, que muestra el pronóstico inmediato para el siguiente año, las líneas del modelo lineal y cuadrático están casi superpuestas, lo cual sugiere que para el próximo año ambos modelos ofrecen proyecciones muy parecidas. Esto podría significar que, a corto plazo, no hay una diferencia significativa entre usar un modelo lineal o cuadrático para predecir las ventas.

En la segunda gráfica, con un horizonte más amplio, las líneas de ambos modelos siguen estando muy juntas en los primeros trimestres después de los datos históricos, lo cual refuerza que las proyecciones a corto plazo son similares. Sin embargo, a medida que el horizonte se extiende, podría empezar a observarse una ligera desviación entre los dos modelos, aunque en esta imagen aún se ven bastante alineados.

En conclusión, en base a la cercanía de las líneas en ambas gráficas, se puede interpretar que ambos modelos están proporcionando predicciones casi idénticas para los próximos trimestres, sin mostrar una ventaja clara de uno sobre el otro en este rango de tiempo. Sin embargo como ya se mencionó anteriormente la contribución del término cuadrático no es estadísticamente significativa por lo que se recomienda utilizar el modelo de tendencia lineal para describir y proyectar la tendencia en las ventas desestacionalizadas, dado que proporciona un ajuste adecuado sin añadir términos innecesarios.