Séries temporelles – Correction TD/TP 2

Manipuler des objets `ts` dans `R`

M1 Actuariat – Université Gustave Eiffel

Cette correction est rédigée comme un rendu possible. Les commentaires peuvent varier, mais les points essentiels sont : vérifier l'horloge, utiliser `window()` sans perdre la structure temporelle, interpréter `time()` et `cycle()`, et traiter les valeurs manquantes explicitement.

Préparation

```
data(co2)
data(UKgas)

x <- co2
gas <- UKgas
```

Exercice 1 – Contrôler l'horloge d'une série

1) Le code demandé est :

```
start(x); end(x); frequency(x)
start(gas); end(gas); frequency(gas)
```

2) On obtient le tableau suivant.

Série	Début	Fin	Fréquence	Interprétation
`co2`	janvier 1959	décembre 1997	12	série mensuelle : 12 observations par année
`UKgas`	1er trimestre 1960	4e trimestre 1986	4	série trimestrielle : 4 observations par année

3) `frequency(x) = 12` signifie que `R` interprète `x` comme une série mensuelle, avec douze observations dans une année. Ce nombre ne mesure pas une périodicité estimée : c'est une information de calendrier donnée à `R`.

Exercice 2 – Chronogramme et zoom avec `window()`

1) Le chronogramme complet est obtenu avec :

```
plot.ts(x, xlab = "annee", ylab = "CO2 (ppm)",
        main = "Concentration mensuelle de CO2")
```

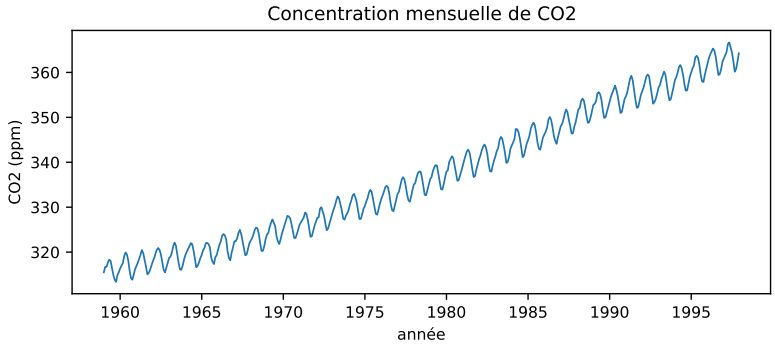


FIGURE 1 – Chronogramme complet de `co2`.

2) La sous-série 1990–1992 se construit avec `window()`.

```
x.9092 <- window(x, start = c(1990, 1), end = c(1992, 12))
plot.ts(x.9092, xlab = "annee", ylab = "CO2 (ppm)",
        main = "CO2 : zoom 1990-1992")
```

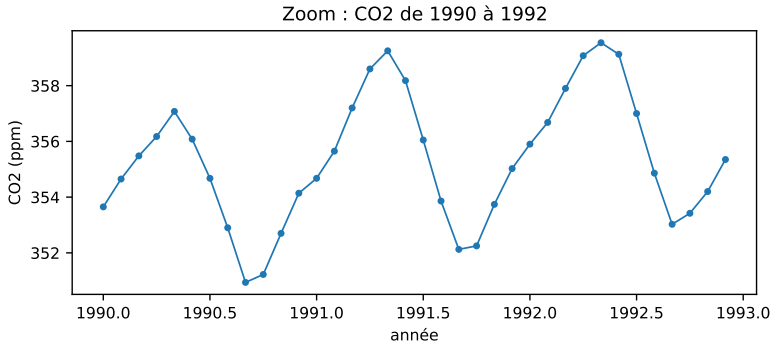


FIGURE 2 – Zoom sur `co2` de janvier 1990 à décembre 1992.

3) Le zoom rend beaucoup plus visible le motif annuel : la série monte et descend à l'intérieur de chaque année. Sur le graphique complet, ce motif existe aussi, mais la tendance croissante domine visuellement.

4) On vérifie :

```
start(x.9092); end(x.9092); frequency(x.9092)
```

On constate que `x.9092` commence en janvier 1990, finit en décembre 1992, et garde la fréquence 12. La fonction `window()` conserve donc l'information temporelle.

Exercice 3 – Temps numérique et cycle annuel

1) Le code est :

```
head(time(x), 12)
head(cycle(x), 12)
```

Les premières valeurs de `time(x)` sont les dates numériques de la première année ; les premières valeurs de `cycle(x)` sont $1, 2, \ldots, 12$.

2) `time(x)` donne la date absolue associée à chaque observation. Pour une série mensuelle, les dates sont codées numériquement à l'intérieur de l'année. `cycle(x)` donne la position dans le cycle annuel : 1 pour janvier, 2 pour février, etc.

3) `cycle(x)` peut servir à comparer les mois entre eux, à construire des indicatrices mensuelles, ou à faire des graphiques saisonniers. C'est une façon de passer de la question « quand ? » à la question « quel mois de l'année ? ».

Exercice 4 – Différence saisonnière

1) La différence saisonnière de retard 12 se calcule ainsi :

```
dx <- diff(x, lag = 12)
plot.ts(dx, xlab = "annee", ylab = "difference annuelle",
        main = "CO2 : difference saisonniere")
abline(h = 0, lty = 2)
```

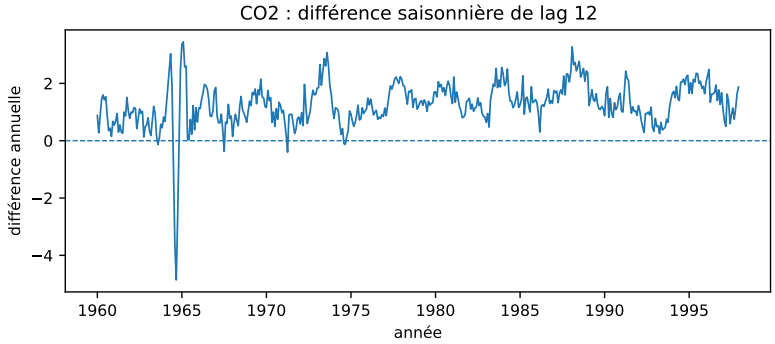


FIGURE 3 – Différence entre chaque mois et le même mois de l'année précédente.

2) La série `dx` représente, pour chaque mois, l'écart entre la concentration de CO2 observée ce mois-là et la concentration observée le même mois l'année précédente. Par exemple, une valeur de 1.5 signifie environ 1.5 ppm de plus que douze mois avant.

3) La différence saisonnière réduit fortement l'oscillation annuelle, car on compare les mêmes mois entre eux. En revanche, elle ne supprime pas toute structure : la série reste plutôt positive et varie au cours du temps. Cela signifie que la concentration augmente encore d'une année sur l'autre, mais à un rythme qui fluctue.

Exercice 5 – Valeurs manquantes

1) On crée la lacune artificielle avec :

```
x.miss <- x
t <- time(x.miss)
ind <- t >= 1991 + 3/12 & t <= 1991 + 8/12
x.miss[ind] <- NA
```

2) Il y a six valeurs manquantes. Le graphique du zoom montre un trou dans la courbe.

```
sum(is.na(x.miss))
plot.ts(window(x.miss, start = c(1990,1), end = c(1992,12)),
        xlab = "annee", ylab = "CO2 (ppm)",
        main = "CO2 avec valeurs manquantes")
```

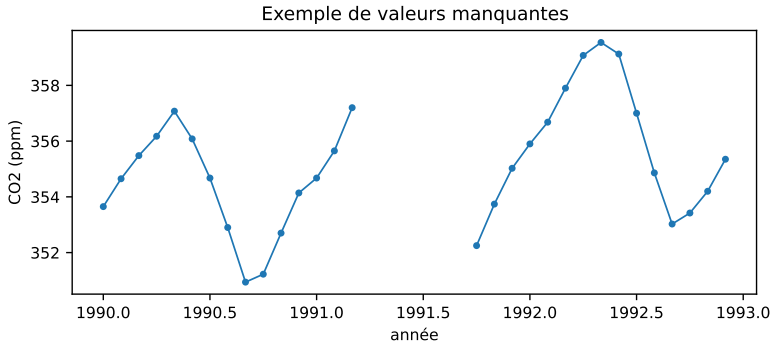


FIGURE 4 – Courte période manquante introduite artificiellement.

3) Une valeur manquante signifie que l'observation n'est pas disponible. Elle ne signifie pas que la concentration vaut zéro. La remplacer par zéro créerait une rupture artificielle énorme et fausserait les graphiques comme les calculs.

4) Une interpolation linéaire possible est :

```
t <- time(x.miss)
ok <- !is.na(x.miss)
x.interp <- x.miss
x.interp[!ok] <- approx(x = t[ok], y = as.numeric(x.miss)[ok],
                        xout = t[!ok])$y

plot.ts(window(x.interp, start = c(1990,1), end = c(1992,12)),
        xlab = "annee", ylab = "CO2 (ppm)",
        main = "CO2 : interpolation lineaire")
```

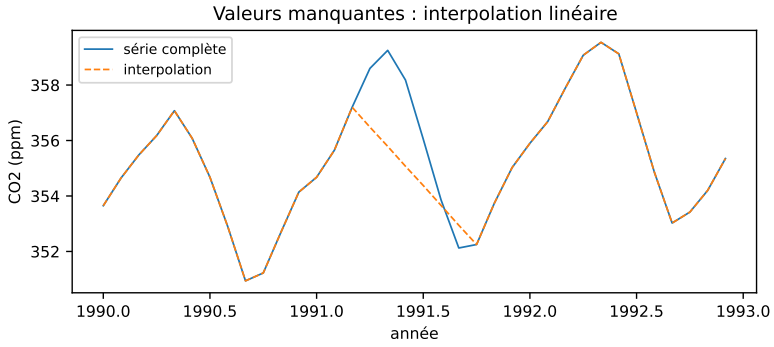


FIGURE 5 – Interpolation linéaire des valeurs manquantes.

Cette interpolation n'est pas une observation réelle. C'est une reconstruction raisonnable pour une courte lacune dans une série régulière. Elle peut être utile pour un graphique ou un calcul exploratoire, mais elle doit être signalée.