MAP 565
Time series analysis : Lecture III

François Roueff
http://perso.telecom-paristech.fr/~roueff/

Telecom ParisTech – École Polytechnique

December 16, 2015

Outline of the course

- ▷ Stochastic modeling
  - I Random processes.
  - II Spectral representation.
- ▷ Linear models
  - III Linear filtering, innovation process. $\leftarrow$
  - IV ARMA processes.
  - V Linear forecasting.
- ▷ Statistical inference
  - VI Overview of goals and methods
  - VII Asymptotic statistics in a dependent context.
- ▷ Non-linear models
  - VIII Standard models for financial time series.
  - IX Complements.

$\leftarrow$: we are here.

Outline of Lecture III

1. Convolution in ℓ^1
   - Basic definitions
   - 2nd order properties
2. Linear filtering in the spectral domain
   - Filtering a white noise
   - The general case
3. Innovations and Wold decomposition
   - Innovation process
   - Complement : Wold decomposition
4. An illustrative example with R

Convolution in ℓ^1

Denote

$$\ell^1 = \left\{ \psi \in \mathbb{C}^{\mathbb{Z}} \;:\; \sum_k |\psi_k| < \infty \right\} .$$

Define the linear filter with impulse response $\psi \in \ell^1$ by the convolution

$$\mathrm{F}_\psi : x = (x_t)_{t\in\mathbb{Z}} \mapsto y = \psi \star x \;, \quad y_t = \sum_{k\in\mathbb{Z}} \psi_k x_{t-k}, \quad t \in \mathbb{Z} \;.$$

Definition : types of filters

- If ψ is finitely supported, F_ψ is called a finite impulse response (FIR) filter.
- If $\psi_t = 0$ for all $t < 0$, F_ψ is said to be causal.
- If $\psi_t = 0$ for all $t \geq 0$, F_ψ is said to be anticausal.

Set of definition

FIR filter

When ψ is finitely supported, we may write

$$\mathrm{F}_{\psi} = \sum_{k \in \mathbb{Z}} \psi_k \, \mathrm{B}^k ,$$

where $\mathrm{B} = \mathrm{S}^{-1}$ is the Backshift operator.

If ψ is not finitely supported, F_{ψ} is well defined only on

$$\ell_{\psi} = \left\{ (x_t)_{t \in \mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}} \ : \ \text{for all } t \in \mathbb{Z}, \sum_{k \in \mathbb{Z}} |\psi_k \, x_{t-k}| < \infty \right\} .$$

Convolution filtering of a time series

Theorem

Let $\psi \in \ell^1$. Then, for all random process $X = (X_t)_{t\in\mathbb{Z}}$ such that

$$\sup_{t\in\mathbb{Z}} \mathbb{E}|X_t| < \infty ,$$

we have $X \in \ell_\psi$ a.s.

If moreover

$$\sup_{t\in\mathbb{Z}} \mathbb{E}\left[|X_t|^2\right] < \infty ,$$

then, for all $t \in \mathbb{Z}$, we have that

$$Y_t = \sum_{k\in\mathbb{Z}} \psi_k X_{t-k} \quad \text{is absolutely convergent in } L^2.$$

Moreover $(Y_t)_{t\in\mathbb{Z}} = \mathrm{F}_\psi(X)$ a.s.

Convolution filtering of weakly stationary time series

Corollary

Let $\psi \in \ell^1$. Then, if $X = (X_t)_{t\in\mathbb{Z}}$ is weakly stationary then $Y = \mathrm{F}_\psi(X)$ is well defined and is an L^2 process.

2nd order properties

Moreover, Y is weakly stationary and, denoting by μ, γ and ν the mean, autocovariance function and spectral measure of X, those of Y are given by

(CF-1) $\mu' = \mu \sum_k \psi_k$,

(CF-2) $\gamma'(\tau) = \sum_{\ell,k} \psi_k \overline{\psi_\ell} \gamma(\tau + \ell - k)$,

(CF-3) $\nu'(\mathrm{d}\lambda) = |\psi^*(\lambda)|^2 \; \nu(\mathrm{d}\lambda)$, with

$$\psi^*(\lambda) = \sum_k \psi_k \mathrm{e}^{-\mathrm{i}\lambda k} \; .$$

A simple case : filtered white noise

Let $(X_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,\sigma^2)$. Then the following assertions are equivalent.

(i) The sum $Y_t = \sum_{k\in\mathbb{Z}} \psi_k X_{t-k}$ converges in in L^2.

(ii) The sequence $(\psi_t)_{t\in\mathbb{Z}} \in \ell^2$.

Convergence in L^2 is sufficient to obtain as for ℓ^1 convolution filtering that

$$Y \text{ is weakly stationary with spectral density } f(\lambda) = \frac{\sigma^2}{2\pi} \left|\psi^*(\lambda)\right|^2 \ ,$$

where ψ^* is the transfer function

$$\psi^*(\lambda) = \sum_{k\in\mathbb{Z}} \psi_k \mathrm{e}^{-\mathrm{i}\lambda k} \ .$$

Hence the condition $\psi \in \ell^1$ is too strong in this case.

Spectral representation of filtered white noise

Note that by construction, the process $(Y_t)_{t\in\mathbb{Z}}$ belongs to $\mathcal{H}_\infty^X$.

Using the spectral representation of X, we have that, for all $t \in \mathbb{Z}$,

$$Y_t = \int \mathrm{e}^{\mathrm{i}\lambda t}\, \psi^*(\lambda)\, \mathrm{d}\widehat{X}(\lambda)\ .$$

Here the unitary property corresponds to Parseval's identity :
$\psi^* : \mathbb{T} \to \mathbb{C}$ is such that

$$\int_{\mathbb{T}} |\psi^*|^2 = 2\pi \sum_{k\in\mathbb{Z}} |\psi_k|^2 < \infty\ .$$

How to generalize this to any process X ?

General linear time-invariant filtering

Let $(X_t)_{t\in\mathbb{Z}}$ be a centered weakly stationary process with an arbitrary spectral measure ν.

We can generalize ℓ^1 convolution filtering by setting

$$Y_t = \lim_{n\to\infty} \sum_{k\in\mathbb{Z}} \psi_{n,k} X_{t-k} \ ,$$

where $(\psi_{n,k})_{k\in\mathbb{Z}}$ has finite support for all n and the limit holds in L^2.

The spectral representation of this limit takes the general form

$$Y_t = \int \mathrm{e}^{\mathrm{i}\lambda t} \, g(\lambda) \, \mathrm{d}\widehat{X}(\lambda) \, , \quad t \in \mathbb{Z} \ ,$$

where $g \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \nu)$. We shall denote

$$Y = \widehat{\mathrm{F}}_g(X) \ .$$

General linear time-invariant filtering (cont.)

Observe that, for all $s, t \in \mathbb{Z}$,

$$\mathrm{Cov}\left(Y_s, Y_t\right) = \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}\lambda(s-t)} \, |g(\lambda)|^2 \, \mathrm{d}\nu(\lambda) \; .$$

Hence $Y = \widehat{\mathrm{F}}_g(X)$ is a centered weakly stationary process and its spectral measure has density $|g|^2$ with respect to ν, the spectral measure of X.

Innovation process

Let $(X_t)_{t\in\mathbb{Z}}$ be a centered weakly stationary process with spectral measure ν. Its linear past up to time t is defined as

$$\mathcal{H}_t^X = \overline{\mathrm{Span}}\,(X_s, s \leq t) \ .$$

Linear prediction and innovation process

- The best linear predictor is defined as the closest element of $\mathcal{H}_{t-1}^X$ to X_t in the L^2 sense and is denoted by

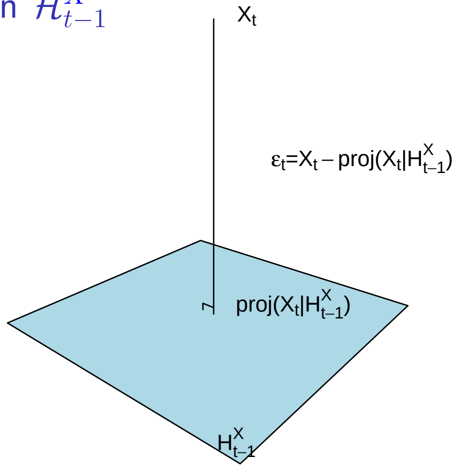
$$\mathrm{proj}\left(X_t\middle|\,\mathcal{H}_{t-1}^X\right) = \underset{Y\in\mathcal{H}_{t-1}^X}{\mathrm{argmin}}\,\mathbb{E}\left[\left|X_t - Y\right|^2\right] \ .$$

- The innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ of X is defined by

$$\epsilon_t = X_t - \mathrm{proj}\left(X_t\middle|\,\mathcal{H}_{t-1}^X\right), \qquad t\in\mathbb{Z} \ .$$

It follows that $(\epsilon_t)_{t\in\mathbb{Z}}$ is a centered orthogonal sequence.

Projection on $\mathcal{H}_{t-1}^{X}$



Since L^2 is a Hilbert space and $\mathcal{H}_{t-1}^{X}$ is a closed linear subspace, they are characterized by the two assertions :

(i) $\operatorname{proj}\left(X_t \middle| \mathcal{H}_{t-1}^{X}\right) \in \mathcal{H}_{t-1}^{X}$

(ii) $\epsilon_t \perp \mathcal{H}_{t-1}^{X}$

Projection in the spectral domain

Since $\mathcal{H}_{t-1}^X \subset \mathcal{H}_\infty^X$, there exists $g_t \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \nu)$ such that

$$\mathrm{proj}\left(X_t | \mathcal{H}_{t-1}^X \right) = \int g_t(\lambda)\ \widehat{X}(\mathrm{d}\lambda) = \widehat{X}(g_t)\ .$$

Moreover, since $\widehat{X}$ is a unitary operator, g_t minimizes

$$g \mapsto \int_{\mathbb{T}} \left| \mathrm{e}^{\mathrm{i}t\lambda} - g(\lambda) \right|^2 \nu(\mathrm{d}\lambda) = \int_{\mathbb{T}} \left| 1 - \mathrm{e}^{-\mathrm{i}t\lambda} g(\lambda) \right|^2 \nu(\mathrm{d}\lambda)$$

over

$$g \in \overline{\mathrm{Span}} \left(\mathrm{e}^{\mathrm{i}s\cdot},\ s < t \right) \iff \mathrm{e}^{-\mathrm{i}t\cdot} g \in \overline{\mathrm{Span}} \left(\mathrm{e}^{\mathrm{i}s\cdot},\ s < 0 \right)\ .$$

We conclude that $\mathrm{e}^{-\mathrm{i}t\cdot} g_t = g_0$, that is,

$$\mathrm{proj}\left(X_t | \mathcal{H}_{t-1}^X \right) = \int \mathrm{e}^{\mathrm{i}t\lambda} g_0(\lambda)\ \widehat{X}(\mathrm{d}\lambda)\ . \tag{1}$$

The innovation process is a white noise

By (1), it follows that the innovation reads in the spectral domain as

$$\epsilon_t = X_t - \mathrm{proj}\left(X_t \middle| \mathcal{H}_{t-1}^X\right) = \int \mathrm{e}^{\mathrm{i}t\lambda}\ (1 - g_0(\lambda))\ \widehat{X}(\mathrm{d}\lambda)\ . \tag{2}$$

Thus, we have $\epsilon = \widehat{\mathrm{F}}_{(1-g_0)}(X)$ and ϵ is a centered white noise.
We denote

$$\sigma^2 = \mathbb{E}\left[|\epsilon_t|^2\right] = \int_{\mathbb{T}} |1 - g_0(\lambda)|^2\ \nu(\mathrm{d}\lambda)\ ,$$

which does not depend on t.

Conclusion

The innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ of X is a centered white noise. Its variance σ^2 is called the innovation variance of X.

Regular/deterministic /purely non-deterministic processes

It may happen that $\sigma = 0$, the constant process providing a trivial example.

Definition : regular/deterministic processes

A centered weakly stationary process is called regular if its innovation variance is positive, $\sigma > 0$. If $\sigma = 0$, we say that it is deterministic.

Purely non-deterministic processes

Note that $\left(\mathcal{H}_t^X\right)_{t\in\mathbb{Z}}$ is an increasing sequence of linear spaces, the union of which has closure $\mathcal{H}_\infty^X$. Let us denote their intersection by $\mathcal{H}_{-\infty}^X$ so that

$$\mathcal{H}_{-\infty}^X \subset \cdots \subset \mathcal{H}_{t-1}^X \subset \mathcal{H}_t^X \subset \mathcal{H}_{t+1}^X \subset \cdots \subset \mathcal{H}_\infty^X \ .$$

- If X is deterministic then $\left(\mathcal{H}_t^X\right)_{t\in\mathbb{Z}}$ is a constant sequence. Then $\mathcal{H}_\infty^X = \mathcal{H}_{-\infty}^X$.
- If X is regular and $\mathcal{H}_{-\infty}^X = \{0\}$, we say that X is purely non-deterministic.

Examples : regular VS deterministic.

- Constant processes are deterministic.
- A centered weakly stationary process X is a white noise if and only if its innovation process $\epsilon = X$.
- Consider the harmonic process X with spectral measure $\nu|_{(-\pi,\pi]} = \sum_{k=1}^{p} \sigma_k^2 \delta_{\lambda_k}$ where $\lambda_1, \ldots, \lambda_p$ are frequencies in $(-\pi, \pi]$ and $\sigma_1^2, \ldots, \sigma_p^2 > 0$. So X has autocovariance function

$$\gamma(\tau) = \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}\tau\lambda}\, \mathrm{d}\nu(\lambda) = \sum_{k=1}^{p} \sigma_k^2 \mathrm{e}^{\mathrm{i}\tau\lambda_k}, \qquad \tau \in \mathbb{Z} .$$

It follows that, for any $n \geq 1$, $\mathrm{Cov}(\begin{bmatrix} X_1 & \ldots & X_n \end{bmatrix}^T)$ has rank at most p and we conclude that X is deterministic.

Examples : purely non-deterministic, or not.

- Let $Z \sim \mathrm{WN}(0, \sigma^2)$ with $\sigma^2 > 0$. Then $\mathcal{H}_{-\infty}^{Z} \perp \mathcal{H}_{\infty}^{Z}$ and $\mathcal{H}_{-\infty}^{Z} \subset \mathcal{H}_{\infty}^{Z}$, so $\mathcal{H}_{-\infty}^{Z} = \{0\}$ and Z is purely non-deterministic.
- Define now $X_t = \sum_{k \geq 0} \psi_k Z_{t-k}$ for some $\psi \in \ell^1$. Then X is a centered weakly stationary process. Note that $\mathcal{H}_t^X \subseteq \mathcal{H}_t^Z$ and thus X is either deterministic or purely non-deterministic.
- Let $Y_t = Z_t + W$ for all $t \in \mathbb{Z}$, where W is a centered L^2 r.v., uncorrelated with Z. Then one can show that $W \in \mathcal{H}_{-\infty}^{Y}$. It follows that $\mathcal{H}_t^Y = \mathrm{Span}\,(W) \overset{\perp}{\oplus} \mathcal{H}_t^Z$, Y has innovation Z and

$$\mathcal{H}_{-\infty}^{Y} = \mathrm{Span}\,(W) \ .$$

It is a regular process but it is not purely non-deterministic.

Conclusion

- The class of centered weakly stationary process is divided into two disjoint subclasses : deterministic processes and regular processes.
- Regular processes can be purely non-deterministic, or not.

Wold decomposition

A regular process can be uniquely decomposed as the sum of a deterministic process and a purely non-deterministic process which are uncorrelated. This is called the Wold decomposition.

Projection on the innovation process

Let X be a centered regular weakly stationary process and let $(\epsilon_t)_{t\in\mathbb{Z}}$ be its innovation process and σ^2 its innovation variance.

Define

$$U_t := \mathrm{proj}\left(X_t \middle| \mathcal{H}_t^{\epsilon}\right) = \sum_{k\geq 0} \psi_k \epsilon_{t-k} \; .$$

where, for all $k = 0, 1, \ldots$

$$\psi_k = \frac{\langle X_t, \epsilon_{t-k}\rangle}{\sigma^2} = \frac{\mathrm{Cov}\left(X_t, \epsilon_{t-k}\right)}{\sigma^2} \; . \tag{3}$$

Using the spectral representation (2), we have

$$\langle X_t, \epsilon_{t-k}\rangle = \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}t\lambda}\overline{\mathrm{e}^{\mathrm{i}(t-k)\lambda}(1 - g_0(\lambda))}\,\nu(\mathrm{d}\lambda) = \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}k\lambda}\overline{(1 - g_0(\lambda))}\,\nu(\mathrm{d}\lambda) \; .$$

Thus we note that ψ_k does not depend on t.

Space decomposition

Recall that

$$X_t = \underbrace{\mathrm{proj}\left(X_t \middle| \mathcal{H}^X_{t-1}\right)}_{\in \mathcal{H}^X_{t-1}} \qquad + \underbrace{\epsilon_t}_{\in \mathrm{Span}(\epsilon_t) \perp \mathcal{H}^X_{t-1}} .$$

In particular we get that $\psi_0 = 1$ and, using an induction on $s < t$,

$$\begin{aligned} \mathcal{H}^X_t &= \mathcal{H}^X_{t-1} \overset{\perp}{\oplus} \mathrm{Span}(\epsilon_t) \\ &= \mathcal{H}^X_s \overset{\perp}{\oplus} \mathrm{Span}(\epsilon_k,\ s < k \leq t) \ . \end{aligned} \tag{4}$$

Then, letting $s \to -\infty$, we get that

$$\mathcal{H}^X_t = \mathcal{H}^X_{-\infty} \overset{\perp}{\oplus} \mathcal{H}^\epsilon_t \ .$$

So we have

$$V_t := \mathrm{proj}\left(X_t \middle| \mathcal{H}^X_{-\infty}\right) = X_t - U_t \ .$$

Wold decomposition

Definition : Wold decomposition

The decomposition $X_t = U_t + V_t$ is called the Wold decomposition.

Theorem : Wold decomposition

Let $(X_t)_{t\in\mathbb{Z}}$ be a centered weakly stationary process. Let $(\epsilon_t)_{t\in\mathbb{Z}}$ be the innovation process and , $(U_t)_{t\in\mathbb{Z}}$ and $(V_t)_{t\in\mathbb{Z}}$ be the two processes resulting in the Wold decomposition $X = U + V$ as above. Then the following facts hold.

- U and V are two uncorrelated processes.
- $(U_t)_{t\in\mathbb{Z}}$ is a regular purely non-deterministic process, $\mathcal{H}_t^U = \mathcal{H}_t^\epsilon$ and U has innovation ϵ.
- V is deterministic and $\mathcal{H}_{-\infty}^V = \mathcal{H}_{-\infty}^X$.

```
############################################
#      Empirical linear prediction         #
############################################
lin_pred <- function(x,p=floor(length(x)/10),l=1)
# learn prediction coefficients
# Input : x : data
#         p : number of prediction coeff.
#         l : lag of prediction
# Output : prediction coeff.
{
  n <- length(x)
  y <- x[(p+l):n]
  X <- x[p:(n-l)]
  if (p>1) {
    for (j in 1:(p-1)){
      X=cbind(X,x[(p-j):(n-j-l)])
    }
  }
  tol = sqrt(.Machine$double.eps)
  a <- t(X) %*% X
  return(solve(a+norm(a)*tol*diag(p) ,t(X) %*% y))
}
comp_lin_pred <- function(x,p=floor(length(x)/10),l=1,
                          predlen=floor(length(x)/10))
# Compute predictions of the last part of x
# Use first part  of x to estimate
# p linear regression coefficients
# Input : x : data
#         p : number of prediction coeff.
#         l : lag of prediction
#         predlen : sample size of the last part to predict
# Output : predicted values
{
  n <- length(x)
  nh <- n-predlen
  # empirical linear fit
```

```
  reg <- lin_pred(x[1:nh],p,l)
  # start prediction
  xp <- numeric()
  for (j in (nh+l):n){
    xp <- c(xp,
            matrix(x[seq(j-l,(j-l-p+1),-1)],nrow=1) %*% reg)
  }
  return(xp)
}
plot_pred <- function(x,xp,dx=0.15,sublen=10,title='',
                      plot_mov=TRUE,plot_all=TRUE)
# Plot successive predictions of the last part of x
# using predictors
# Input : x : data
#         xp : predictors
#         dx : sleeping time netween successive plots
#         sublen : plotted subsample size
#         title : Plot main title
#         plot_mov : if TRUE, plot moving predictions
#         plot_all : if TRUE, plot all predictions
{
  n <- length(x)
  nh <- n-length(xp)
  # set y range
  absymax <- max(c(abs(min(x)-abs(min(x)/50)),
                  abs(max(x)+abs(max(x)/50))))
  xminmax <- c(-absymax,absymax)
  if (plot_mov) {
    # moving plots of x (black) and xp(red)
    for (j in (nh+sublen):n)
      {
        tt <- (j-sublen+1):j
        plot(tt, x[tt],type='o',ylim=xminmax,ylab='',
             xlab='time',main=title)
        lines(tt, xp[tt-nh],type='o',col=2,pch=3)
        legend("topleft",legend = c("obs","pred"),
```

```
              lty=rep(1,2),pch=c(1,3),col = c(1,2))
        Sys.sleep(dx)
      }
  }
  if (plot_all) {
    #plots of x (black) and xp (red) and squared errors
    tt <- (nh+1):n
    op <- par(mfrow=c(2,1))
    plot(tt, x[tt],type='o',ylim=xminmax,ylab='',
         xlab='time',main=title)
    lines(tt, xp[tt-nh],type='o',col=2,pch=3)
    legend("topleft",legend = c("obs","pred"),
           lty=rep(1,2),pch=c(1,3),col = c(1,2))
    plot(tt, (x[tt])**2,type='o',ylim=c(0,absymax**2),
         ylab='', xlab='time')
    lines(tt, ((x[tt]-xp[tt-nh]))**2,ylab='',
          type='l',xlab='time',col=3)
    legend("topleft",legend = c("Sq. obs","Sq. Err."),
           lty=rep(1,2),pch=c(1,46),col = c(1,3))
    par(op)
  }
}
plot_lin_pred <- function(x,p=floor(length(x)/10),l=1,
                          predlen=floor(length(x)/10),
                          dx=0.15, sublen=10,title='',
                          plot_mov=TRUE,plot_all=TRUE)
# Plot successive predictions of the last part of x
# Use first part  of x to estimate
# p linear regression coefficients
# Input : x : data
#         p : number of prediction coeff.
#         l : lag of prediction
#         predlen : sample size of the last part to predict
#         dx : sleeping time netween successive plots
#         sublent : plotted subsample size
#         title : Plot main title
```

```
#         plot_mov : if TRUE, plot moving predictions
#         plot_all : if TRUE, plot all predictions
{
  xp <- comp_lin_pred(x,p,l,predlen)
  plot_pred(x,xp,dx,sublen,title,plot_mov,plot_all)
}
############################################
#      Application to White noise          #
############################################
#generate white noise
# length of the time series
n <- 2^10
Z <- rnorm(n,mean=0,sd=1)
# linear prediction
plot_lin_pred(x=Z,p=2^3,title='Prediction of white noise')
# try a longer time series
n <- 2^20
Z <- rnorm(n,mean=0,sd=1)
# linear prediction
plot_lin_pred(x=Z,p=2^3,predlen=100,
              title='Prediction using longer sample')
############################################
# Harmonic proc. has zero innovation       #
############################################
# number of frequencies
nfreq <- 2^4
# length of the time series
n <- 2^10
# random generators
# random phases and  amplititudes
set.seed(1)
phase <- runif(nfreq)*2*pi; amp <- rnorm(n)
# set of frequencies picked randomly
lam <- runif(nfreq)*pi
# grenerate signal by adding frequencies
tt <- 1:n; S <- 0
```

```
for (j in 1:nfreq){
  S <- S+amp[j]*cos(lam[j]*tt+phase[j])
}
S <- S/sqrt(var(S))
#plot the begining of signal
ts.plot(S[1:2^7])
# linear prediction
plot_lin_pred(x=S,p=2^3,
              title='Prediction of harmonic process')
# linear prediction with a longer past
plot_lin_pred(x=S,p=2^5,
              title='Prediction using longer past')
############################################
#         A more interesting one           #
############################################
unif_gen_recursive <- function(n=2^8){
  x <- 2*runif(1)
  for (t in 2:n){
    x <- c(x,x[t-1]*0.5+ (runif(1)>0.5))
  }
  return(x-1)
}
# geneate AR(1) uniform process
# length of the time series
n <- 2^10
X <- unif_gen_recursive(n)
ts.plot(X[1:2^7])
# linear prediction
plot_lin_pred(x=X,p=1,
              title='Prediction of AR(1) process')
# add an harmonic process
ts.plot(X+S)
# and predict again
plot_lin_pred(x=X+S,p=2^5,
              title='AR(1) + harmonic process')
```

```
# try to predict further ahead
l <- 2^4
graphics.off()
plot_lin_pred(x=S,p=2^5,l,
              title=paste('Prediction of harmonic process',
                l,' samples ahead',sep=''),plot_mov=FALSE)
x11()
plot_lin_pred(x=X,p=1,l,
                title=paste('Prediction of AR(1) ',
                  l,' samples ahead',sep=''),plot_mov=FALSE)
x11()
plot_lin_pred(x=X+S,p=2^5,l,
              title=paste('Prediction of AR(1)+ harmonic ',
                l,' samples ahead',sep=''),plot_mov=FALSE)
```