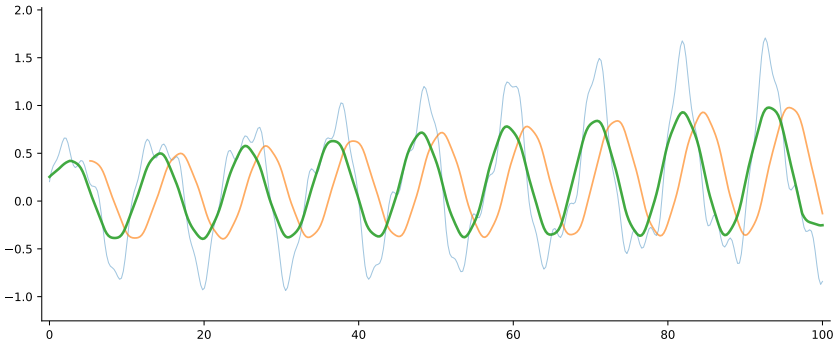
Séries temporelles

Cours 4 – Modèles MA, AR et ARMA stationnaires



M1 Actuariat – Université Gustave Eiffel

Solesne Bourguin – 2026–2027

Fil conducteur de la séance

On a appris à manipuler des transformations linéaires d'une série. On passe maintenant aux premiers **modèles probabilistes stationnaires**.

Fil conducteur de la séance

On a appris à manipuler des transformations linéaires d'une série. On passe maintenant aux premiers **modèles probabilistes stationnaires**.

1. Rappels sur l'idée de filtre, le bruit blanc et l'autocorrélation.
2. Modèles MA (*Moving Average*) : combinaison finie de chocs récents.
3. Modèles AR (*AutoRegressive*) : propagation récursive des valeurs passées.
4. Modèles ARMA (*AutoRegressive Moving Average*) : combinaison des deux mécanismes.
5. Lecture des signatures avec l'autocorrélation (ACF) et l'autocorrélation partielle (PACF).

Fil conducteur de la séance

On a appris à manipuler des transformations linéaires d'une série. On passe maintenant aux premiers **modèles probabilistes stationnaires**.

1. Rappels sur l'idée de filtre, le bruit blanc et l'autocorrélation.
2. Modèles MA (*Moving Average*) : combinaison finie de chocs récents.
3. Modèles AR (*AutoRegressive*) : propagation récursive des valeurs passées.
4. Modèles ARMA (*AutoRegressive Moving Average*) : combinaison des deux mécanismes.
5. Lecture des signatures avec l'autocorrélation (ACF) et l'autocorrélation partielle (PACF).

Objectif : comprendre ce que les modèles MA, AR et ARMA disent sur la **mémoire** d'une série stationnaire.

Rappel 1/3 : l'idée de filtre

Un filtre linéaire remplace X_t par une combinaison de valeurs voisines.

$$(a(B)X)_t = \sum_j a_j X_{t-j}.$$

Filtre	**Écriture**	**Effet**
Différence	$1 - B$	$X_t - X_{t-1}$
Différence saisonnière	$1 - B^s$	$X_t - X_{t-s}$
Moyenne mobile passée	$\frac{1}{3}(1 + B + B^2)$	$\frac{1}{3}(X_t + X_{t-1} + X_{t-2})$

Le filtre décrit explicitement quels retards sont utilisés et avec quels poids.

Rappel 2/3 : bruit blanc et série filtrée

Un bruit blanc centré $Z = (Z_t)$ est une suite de chocs non corrélés :

$$\mathbb{E}(Z_t) = 0, \qquad \mathrm{Cov}(Z_t, Z_{t+h}) = \sigma^2 \mathbf{1}_{\{h=0\}}.$$

Bruit blanc

- pas de dépendance linéaire ;
- pas de mémoire ;
- ACF nulle après le retard 0.

Bruit blanc filtré

- combinaison de chocs ;
- dépendance créée par les chocs communs ;
- mémoire contrôlée par les coefficients.

Les modèles de cette séance construisent des séries dépendantes à partir d'innovations Z_t.

Rappel 3/3 : autocorrélation et autocorrélation partielle

Pour une série stationnaire du second ordre, l'**ACF** signifie *AutoCorrelation Function*, c'est-à-dire la fonction d'autocorrélation :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \qquad \gamma(h) = \mathrm{Cov}(X_t, X_{t+h}).$$

ACF : autocorrélation

- dépendance linéaire totale au retard h ;
- mesure si X_t et X_{t+h} se ressemblent.

PACF : autocorrélation partielle

- dépendance directe au retard h ;
- après retrait des retards intermédiaires.

La PACF signifie *Partial AutoCorrelation Function*. ACF et PACF donnent une première lecture de la mémoire d'une série.

Première partie

Pourquoi introduire les modèles MA, AR et ARMA ?

Après les transformations, il reste souvent une série stationnaire mais dépendante.

La situation visée

On ne modélise pas directement une série qui a encore une tendance forte ou une saisonnalité évidente.

Étape	**Question**	**Outils**
Préparer la série	tendance ? saisonnalité ? échelle ?	log, différences, désaisonnalisation
Modéliser le reste	dépendance temporelle ? mémoire ?	MA, AR, ARMA
Diagnostiquer	le modèle suffit-il ?	résidus, ACF, tests

Les modèles ARMA décrivent la dépendance d'une série que l'on suppose déjà à peu près stationnaire.

Trois mécanismes de mémoire

Les modèles MA, AR et ARMA ne décrivent pas la dépendance de la même manière.

MA

La série est une combinaison finie de chocs récents.

$$X_t = Z_t + \theta Z_{t-1}$$

Mémoire finie des innovations.

AR

La série dépend de ses propres valeurs passées.

$$X_t = \phi X_{t-1} + Z_t$$

Propagation récursive.

ARMA

On combine les deux mécanismes.

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}$$

Mémoire plus flexible.

Notation commune

Dans toute la séance, $Z = (Z_t)$ désigne un bruit blanc centré de variance σ^2.

Les Z_t sont appelés **innovations** : au temps t, Z_t représente la partie nouvelle qui n'est pas expliquée par le passé du modèle.

On utilisera aussi les polynômes en l'opérateur retard B :

$$\Theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q, \qquad \Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p.$$

Le signe moins dans $\Phi(B)$ est une convention classique : elle permet d'écrire les AR sous la forme $\Phi(B)X_t = Z_t$.

Deuxième partie

Modèles MA : moyenne mobile de chocs

Un MA est le modèle le plus proche des filtres finis vus précédemment.

Définition d'un MA(q)

MA signifie *Moving Average*, que l'on traduit par **moyenne mobile**. Ici, « moyenne » est historique : il s'agit surtout d'une **combinaison linéaire mobile** de chocs.

Un processus $X = (X_t)$ est un MA(q) s'il s'écrit

$$X_t = m + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}.$$

En posant $\theta_0 = 1$,

$$X_t = m + \sum_{j=0}^{q} \theta_j Z_{t-j} = m + \Theta(B) Z_t.$$

Les coefficients ne sont pas nécessairement positifs et ne somment pas nécessairement à 1. Ce n'est donc pas toujours une moyenne au sens usuel.

MA(1) : le premier exemple

Considérons

$$X_t = Z_t + \theta Z_{t-1}.$$

	X_{t-1}	X_t	X_{t+1}
Z_{t-1}	présent	présent	absent
Z_t	absent	présent	présent
Z_{t+1}	absent	absent	présent

Deux observations voisines partagent un choc commun. Deux observations séparées d'au moins deux dates ne partagent plus aucun choc. C'est la mémoire finie du MA(1).

MA(1) : moyenne et variance

Avec $X_t = Z_t + \theta Z_{t-1}$ et $\mathbb{E}(Z_t) = 0$,

$$\mathbb{E}(X_t) = \mathbb{E}(Z_t) + \theta \mathbb{E}(Z_{t-1}) = 0.$$

MA(1) : moyenne et variance

Avec $X_t = Z_t + \theta Z_{t-1}$ et $\mathbb{E}(Z_t) = 0$,

$$\mathbb{E}(X_t) = \mathbb{E}(Z_t) + \theta \mathbb{E}(Z_{t-1}) = 0.$$

Pour la variance, les deux termes sont non corrélés :

$$\gamma(0) = \mathrm{Var}(X_t) = \mathrm{Var}(Z_t) + \theta^2 \mathrm{Var}(Z_{t-1}) = \sigma^2(1 + \theta^2).$$

MA(1) : moyenne et variance

Avec $X_t = Z_t + \theta Z_{t-1}$ et $\mathbb{E}(Z_t) = 0$,

$$\mathbb{E}(X_t) = \mathbb{E}(Z_t) + \theta \mathbb{E}(Z_{t-1}) = 0.$$

Pour la variance, les deux termes sont non corrélés :

$$\gamma(0) = \mathrm{Var}(X_t) = \mathrm{Var}(Z_t) + \theta^2 \mathrm{Var}(Z_{t-1}) = \sigma^2(1 + \theta^2).$$

Règle importante : dans une covariance, les termes contenant deux innovations différentes disparaissent. Il ne reste que les termes où la *même* innovation est présente dans les deux variables comparées.

MA(1) : covariance au retard 1

On calcule $\gamma(1) = \mathrm{Cov}(X_t, X_{t+1})$.

$$X_t = Z_t + \theta Z_{t-1}, \qquad X_{t+1} = Z_{t+1} + \theta Z_t.$$

$$\begin{aligned}\gamma(1) &= \mathrm{Cov}(Z_t + \theta Z_{t-1}, Z_{t+1} + \theta Z_t) \\ &= \theta \,\mathrm{Var}(Z_t) \\ &= \theta\sigma^2.\end{aligned}$$

Le seul choc commun entre X_t et X_{t+1} est Z_t.

MA(1) : coupure de l'ACF

Pour $|h| \geq 2$, X_t et X_{t+h} ne partagent aucun choc commun.

$$\gamma(h) = 0, \qquad |h| \geq 2.$$

Ainsi,

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta}{1+\theta^2}, \qquad \rho(h) = 0 \quad \text{si } |h| \geq 2.$$

Signature d'un MA(1) : l'ACF théorique est coupée après le retard 1.

MA(q) : formule générale

Pour

$$X_t = \mu + \sum_{j=0}^{q} \theta_j Z_{t-j}, \qquad \theta_0 = 1,$$

le processus est stationnaire du second ordre et

$$\gamma(h) = \begin{cases} \sigma^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \leq h \leq q, \\ 0, & h > q. \end{cases}$$

MA(q) : formule générale

Pour

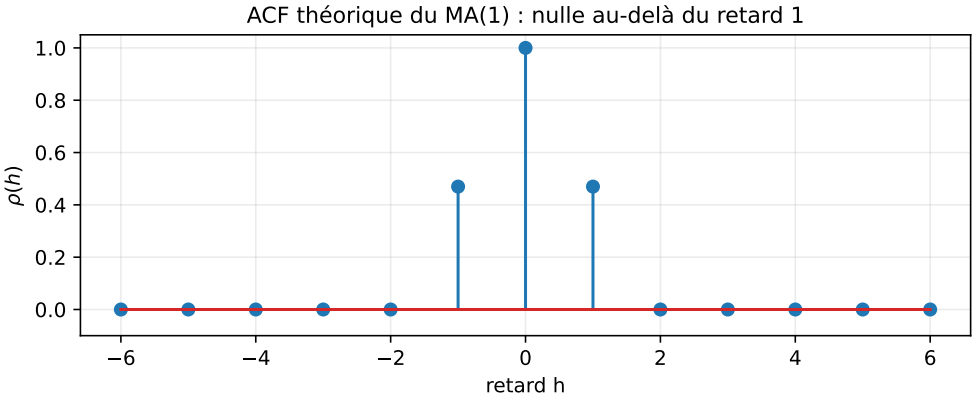
$$X_t = \mu + \sum_{j=0}^{q} \theta_j Z_{t-j}, \qquad \theta_0 = 1,$$

le processus est stationnaire du second ordre et

$$\gamma(h) = \begin{cases} \sigma^2 \displaystyle\sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \leq h \leq q, \\ 0, & h > q. \end{cases}$$

Un MA(q) a une mémoire de longueur q : au-delà du retard q, les observations ne partagent plus d'innovation commune.

Visualiser la signature MA



Pour un MA(1) théorique, seule la barre au retard 1 peut être non nulle après le retard 0.

Inversibilité

Dans un MA(1),

$$\rho(1) = \frac{\theta}{1+\theta^2}.$$

Mais

$$\frac{\theta}{1+\theta^2} = \frac{1/\theta}{1+1/\theta^2} \qquad (\theta \neq 0).$$

Inversibilité

Dans un MA(1),

$$\rho(1) = \frac{\theta}{1+\theta^2}.$$

Mais

$$\frac{\theta}{1+\theta^2} = \frac{1/\theta}{1+1/\theta^2} \qquad (\theta \neq 0).$$

Deux paramètres différents, θ et $1/\theta$, peuvent donner la même ACF. Le même comportement de covariance peut donc avoir deux écritures MA différentes.

On choisit comme écriture canonique l'écriture **inversible** : les innovations Z_t doivent pouvoir être reconstruites à partir du présent et du passé observé de la série.

Inversibilité du MA(1)

Si la série est centrée et

$$X_t = (1 + \theta B) Z_t,$$

alors reconstruire les innovations revient à inverser le filtre $1 + \theta B$. Lorsque $|\theta| < 1$,

$$(1 + \theta B)^{-1} = 1 - \theta B + \theta^2 B^2 - \theta^3 B^3 + \cdots .$$

Donc

$$Z_t = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \theta^3 X_{t-3} + \cdots .$$

Pour un MA(1), l'écriture canonique est celle avec $|\theta| < 1$, c'est-à-dire celle dont la racine de $\Theta(z) = 1 + \theta z$ est hors du disque unité.

Troisième partie

Modèles AR : propagation récursive

Dans un AR, la valeur présente dépend directement des valeurs passées de la série.

Définition d'un AR(p)

AR signifie *AutoRegressive*, ou **autorégressif** : on explique X_t par ses propres valeurs passées.

On écrit un AR(p) autour de sa moyenne m sous la forme

$$X_t - m = \phi_1(X_{t-1} - m) + \cdots + \phi_p(X_{t-p} - m) + Z_t.$$

Sous forme polynomiale,

$$\Phi(B)(X_t - m) = Z_t, \qquad \Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p.$$

Ici m désigne directement la moyenne de la série. Si $\mathbb{E}[Z_t] = 0$, alors en prenant l'espérance dans l'équation, on obtient bien $\mathbb{E}[X_t] = m$.

AR(1) : équation et mémoire

Le modèle AR(1) centré est

$$X_t = \phi X_{t-1} + Z_t.$$

Si $\phi > 0$, une valeur positive tend à être suivie par une valeur positive. La série est persistante.

Si $\phi < 0$, une valeur positive tend à être suivie par une valeur négative. La série alterne davantage.

Un choc ne disparaît pas après une seule date : il se propage par la récurrence.

AR(1) comme filtre infini

On remplace récursivement X_{t-1} :

$$\begin{aligned} X_t &= \phi X_{t-1} + Z_t \\ &= \phi(\phi X_{t-2} + Z_{t-1}) + Z_t \\ &= Z_t + \phi Z_{t-1} + \phi^2 X_{t-2}. \end{aligned}$$

En poursuivant,

$$X_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \cdots,$$

à condition que $|\phi| < 1$.

Un AR(1) causal est un filtre infini du bruit blanc, avec des poids géométriques.

AR(1) : autocovariance et ACF

Si $|\phi| < 1$, alors

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}.$$

On obtient

$$\gamma(h) = \frac{\sigma^2 \phi^{|h|}}{1-\phi^2}, \qquad \rho(h) = \phi^{|h|}.$$

AR(1) : autocovariance et ACF

Si $|\phi| < 1$, alors

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}.$$

On obtient

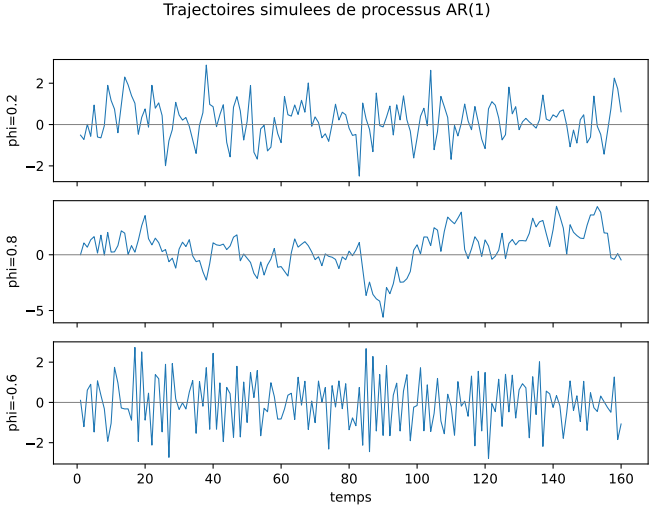
$$\gamma(h) = \frac{\sigma^2 \phi^{|h|}}{1-\phi^2}, \qquad \rho(h) = \phi^{|h|}.$$

$0 < \phi < 1$: ACF positive et décroissante.

$-1 < \phi < 0$: ACF qui alterne de signe.

Contrairement au MA(1), l'ACF d'un AR(1) ne se coupe pas : elle décroît progressivement.

AR(1) : trajectoires simulées



Persistance pour $\phi > 0$; alternance plus marquée pour $\phi < 0$.

Condition de causalité

Pour un AR(p),

$$\Phi(B)X_t = Z_t, \qquad \Phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p.$$

Le modèle AR(p) est causal si les racines de $\Phi(z)$ sont toutes de module strictement supérieur à 1.

Condition de causalité

Pour un AR(p),

$$\Phi(B)X_t = Z_t, \qquad \Phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p.$$

Le modèle AR(p) est causal si les racines de $\Phi(z)$ sont toutes de module strictement supérieur à 1.

Cette condition garantit que $1/\Phi(B)$ se développe en filtre causal infini stable :

$$X_t = \psi(B)Z_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Exemple : AR(1) et racine

Pour un AR(1),

$$\Phi(z) = 1 - \phi z.$$

La racine est

$$z = \frac{1}{\phi}.$$

La condition « racine hors du disque unité » donne

$$\left|\frac{1}{\phi}\right| > 1 \qquad \Longleftrightarrow \qquad |\phi| < 1.$$

On retrouve exactement la condition vue avec la série géométrique.

Fractions rationnelles

Pour un AR centré,

$$\Phi(B)X_t = Z_t.$$

Cette équation ne donne pas encore X_t directement comme fonction des chocs. Pour comprendre la mémoire du modèle, on cherche une écriture causale

$$X_t = \psi(B)Z_t = \psi_0 Z_t + \psi_1 Z_{t-1} + \psi_2 Z_{t-2} + \cdots .$$

En remplaçant dans l'équation AR, il faut donc

$$\Phi(B)\psi(B) = 1, \qquad \text{donc} \qquad \psi(B) = \frac{1}{\Phi(B)}.$$

On obtient une fraction rationnelle parce qu'on inverse un polynôme en B.

Méthode : identifier les coefficients

On travaille d'abord formellement avec une variable z :

$$\Phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \qquad \psi(z) = \sum_{j \geq 0} \psi_j z^j.$$

La relation $\Phi(z)\psi(z) = 1$ donne

$$(1 - \phi_1 z - \cdots - \phi_p z^p)(\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots) = 1.$$

On développe le produit, puis on identifie les coefficients de $z^0, z^1, z^2, \ldots$.

$$\psi_0 = 1, \qquad \psi_j = \phi_1 \psi_{j-1} + \cdots + \phi_p \psi_{j-p}, \quad j \geq 1,$$

avec la convention $\psi_k = 0$ si $k < 0$.

Exemple : développement d'un AR(2)

Supposons

$$\Phi(z) = 1 - 1.4z + 0.45z^2.$$

On cherche $\psi(z) = \psi_0 + \psi_1 z + \psi_2 z^2 + \cdots$ tel que

$$(1 - 1.4z + 0.45z^2)\psi(z) = 1.$$

$$\begin{aligned}
\text{coef. } z^0 : &\quad \psi_0 = 1, \\
\text{coef. } z^1 : &\quad \psi_1 - 1.4\psi_0 = 0, \\
\text{coef. } z^j : &\quad \psi_j - 1.4\psi_{j-1} + 0.45\psi_{j-2} = 0 \quad (j \geq 2).
\end{aligned}$$

Donc

$$\psi_j = 1.4\psi_{j-1} - 0.45\psi_{j-2}.$$

Exemple : premiers coefficients

À partir de

$$\psi_0 = 1, \qquad \psi_j = 1.4\psi_{j-1} - 0.45\psi_{j-2},$$

on obtient

$$\psi_1 = 1.4, \qquad \psi_2 = 1.4^2 - 0.45 = 1.51,$$

$$\psi_3 = 1.4 \times 1.51 - 0.45 \times 1.4 = 1.484, \quad \ldots$$

La causalité garantit que cette suite de coefficients définit bien un filtre infini stable. Ici les racines de $1 - 1.4z + 0.45z^2$ sont hors du disque unité.

Quatrième partie

Modèles ARMA : combiner AR et MA

Un ARMA combine une partie récursive et une partie moyenne mobile.

Définition d'un ARMA(p, q)

ARMA signifie *AutoRegressive Moving Average* : on combine une partie autorégressive et une partie moyenne mobile.

Un processus $X = (X_t)$ est un ARMA(p, q) si

$$\Phi(B)(X_t - m) = \Theta(B)Z_t,$$

avec

$$\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p, \qquad \Theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q.$$

Exemple ARMA$(1, 1)$ centré :

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}.$$

La partie AR propage les chocs ; la partie MA modifie leur impact immédiat.

ARMA comme filtre rationnel

Partons de

$$\Phi(B)(X_t - m) = \Theta(B)Z_t.$$

Si la partie AR est causale, on peut inverser $\Phi(B)$ et écrire

$$X_t - m = \frac{\Theta(B)}{\Phi(B)} Z_t = \psi(B)Z_t.$$

Numérateur Θ

- effet MA ;
- agit sur quelques chocs récents.

Dénominateur Φ

- effet AR ;
- crée une propagation dans le temps.

La fraction rationnelle sert à obtenir les coefficients ψ_j du filtre infini équivalent.

Méthode générale pour un ARMA

On pose

$$\psi(z) = \sum_{j \geq 0} \psi_j z^j, \qquad \Theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q.$$

La relation à résoudre est maintenant

$$\Phi(z)\psi(z) = \Theta(z).$$

En identifiant les coefficients, on obtient

$$\psi_j - \phi_1 \psi_{j-1} - \cdots - \phi_p \psi_{j-p} = \theta_j,$$

avec

$$\theta_0 = 1, \qquad \theta_j = 0 \text{ si } j > q, \qquad \psi_j = 0 \text{ si } j < 0.$$

Même méthode que pour un AR, mais le membre de droite n'est plus seulement 1 : il contient la partie MA.

Exemple : ARMA$(1,1)$

Pour

$$(1-\phi B)(X_t - m) = (1+\theta B)Z_t,$$

on cherche

$$X_t - m = \psi(B)Z_t, \qquad \text{avec} \qquad (1-\phi z)\psi(z) = 1+\theta z.$$

Identification des coefficients :

$$\psi_0 = 1, \qquad \psi_1 - \phi\psi_0 = \theta, \qquad \psi_j - \phi\psi_{j-1} = 0 \quad (j \geq 2).$$

Donc

$$\psi_1 = \phi + \theta, \qquad \psi_j = \phi^{j-1}(\phi+\theta) \quad (j \geq 1).$$

La partie MA modifie les premiers coefficients ; la partie AR impose ensuite une décroissance géométrique.

Causalité et inversibilité

Causalité
Racines de $\Phi(z)$ hors du disque unité.

$$X_t - m \text{ dépend de } Z_t, Z_{t-1}, \ldots$$

La dynamique est stable.

Inversibilité
Racines de $\Theta(z)$ hors du disque unité.

$$Z_t \text{ se reconstruit depuis } X_t, X_{t-1}, \ldots$$

L'écriture des innovations est canonique.

Les racines de Φ contrôlent la stabilité de la dynamique ; les racines de Θ contrôlent le choix d'une représentation unique des innovations.

Pourquoi combiner AR et MA ?

MA seul
Très flexible aux petits retards, mais mémoire strictement finie.

AR seul
Mémoire persistante avec peu de paramètres, mais forme parfois trop contrainte.

ARMA
Permet une dépendance courte flexible et une décroissance progressive ensuite.

Un ARMA$(1,1)$ peut déjà produire des formes d'ACF plus variées qu'un AR(1) ou qu'un MA(1) isolé.

Cinquième partie

ACF, PACF et identification

On lit la mémoire d'une série avec deux graphiques complémentaires.

ACF théorique et ACF empirique

L'ACF théorique d'un processus stationnaire est

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \qquad \gamma(h) = \mathrm{Cov}(X_t, X_{t+h}).$$

À partir d'une série observée $x_1, \ldots, x_n$, on remplace les quantités théoriques par des quantités calculées sur les données.

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t, \qquad \widehat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x}), \qquad \widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}.$$

Le graphique ACF représente les valeurs $\widehat{\rho}(h)$ pour plusieurs retards h.

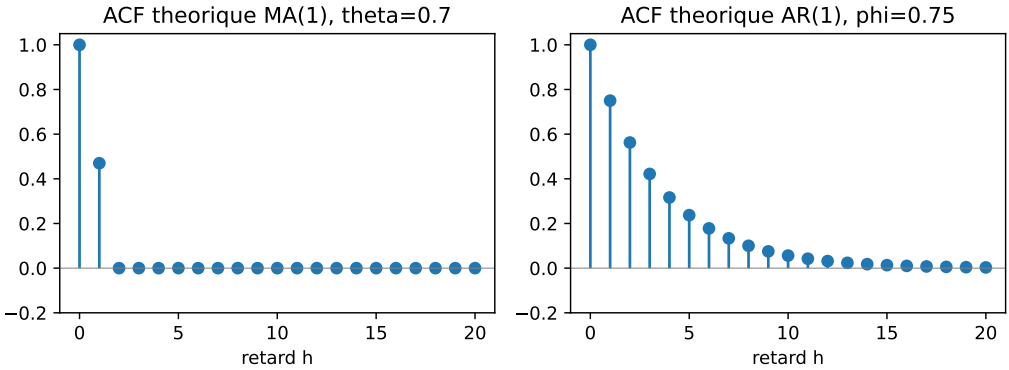
Comment lire une ACF empirique ?

Une grande barre au retard h signifie que x_t et x_{t+h} se ressemblent souvent dans l'échantillon.

Une barre proche de 0 signifie qu'on ne voit pas de corrélation linéaire nette à ce retard.

Une ACF empirique est seulement une estimation de l'ACF théorique : elle fluctue d'un échantillon à l'autre. Les petites barres isolées ne doivent donc pas être surinterprétées.

Formes typiques de l'ACF



Un MA a une ACF coupée ; un AR a une ACF qui décroît progressivement.

Pourquoi l'ACF ne suffit pas

Dans un AR(1),

$$X_t = \phi X_{t-1} + Z_t.$$

Le modèle ne dépend directement que de X_{t-1}. Pourtant,

$$\rho(h) = \phi^{|h|}$$

n'est pas nul pour $h = 2, 3, 4, \ldots$

La dépendance à long retard peut être indirecte : X_t dépend de X_{t-1}, qui dépend de X_{t-2}, etc.

On introduit donc la PACF pour mesurer la dépendance directe au retard h, une fois les retards intermédiaires pris en compte.

PACF théorique : que retire-t-on ?

On note $\alpha(h)$ la **fonction d'autocorrélation partielle** au retard h. On travaille avec la série centrée

$$Y_t = X_t - m.$$

Pour mesurer le lien *direct* entre Y_t et Y_{t-h}, on veut enlever ce qui est déjà expliqué linéairement par les retards intermédiaires

$$Y_{t-1}, Y_{t-2}, \ldots, Y_{t-h+1}.$$

On introduit donc l'espace linéaire

$$\mathcal{H}_h = \text{Vect}\left(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-h+1}\right).$$

La meilleure approximation linéaire de Y_t par les retards intermédiaires est la variable

$$\widehat{Y}_t = a_1 Y_{t-1} + \cdots + a_{h-1} Y_{t-h+1}$$

qui minimise

$$\mathbb{E}\left[(Y_t - Z)^2\right] \qquad \text{parmi tous les } Z \in \mathcal{H}_h.$$

Autrement dit : on choisit les coefficients qui donnent la plus petite erreur quadratique moyenne.

PACF théorique : définition générale

On retire la partie de Y_t expliquée par les retards intermédiaires, et on fait la même chose pour Y_{t-h} :

$$U_t = Y_t - \widehat{Y}_t, \qquad V_t = Y_{t-h} - \widehat{Y}_{t-h},$$

où

$$\widehat{Y}_t = \underset{Z \in \mathcal{H}_h}{\operatorname{argmin}} \, \mathbb{E}\big[(Y_t - Z)^2\big], \qquad \widehat{Y}_{t-h} = \underset{Z \in \mathcal{H}_h}{\operatorname{argmin}} \, \mathbb{E}\big[(Y_{t-h} - Z)^2\big].$$

$$\begin{array}{ccccc} Y_t & \longrightarrow & \text{retirer ce qui vient de } \mathcal{H}_h & \longrightarrow & U_t \\ & & Y_{t-1}, \ldots, Y_{t-h+1} & & \\ Y_{t-h} & \longrightarrow & \text{retirer ce qui vient de } \mathcal{H}_h & \longrightarrow & V_t \end{array}$$

La PACF au retard h est la corrélation entre les deux restes :

$$\alpha(h) = \mathrm{Corr}(U_t, V_t).$$

Pour $h = 1$, il n'y a aucun retard intermédiaire à retirer, donc $\alpha(1) = \rho(1)$.

PACF théorique : autre écriture utile

Une définition équivalente consiste à regarder la meilleure prédiction linéaire de $Y_t = X_t - m$ par ses h derniers retards centrés :

$$Y_t \approx a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_h Y_{t-h}.$$

Les coefficients théoriques sont ceux qui minimisent l'erreur quadratique moyenne :

$$(a_1, \ldots, a_h) = \arg \min_{b_1, \ldots, b_h} \mathbb{E}\left[\left(Y_t - \sum_{j=1}^{h} b_j Y_{t-j}\right)^2\right].$$

La PACF au retard h est le dernier coefficient de cette régression linéaire théorique :

$$\alpha(h) = a_h.$$

L'ACF mesure la dépendance totale au retard h ; la PACF mesure ce que le retard h ajoute après les retards $1, \ldots, h-1$.

PACF empirique : comment la construire ?

À partir d'une série observée $x_1, \ldots, x_n$, on centre d'abord les données :

$$y_t = x_t - \bar{x}.$$

Pour un retard h, on ajuste ensuite la régression linéaire

$$y_t \approx \widehat{a}_1 y_{t-1} + \widehat{a}_2 y_{t-2} + \cdots + \widehat{a}_h y_{t-h}, \qquad t = h+1, \ldots, n.$$

On choisit les coefficients par moindres carrés, c'est-à-dire en minimisant

$$\sum_{t=h+1}^{n} \left(y_t - \sum_{j=1}^{h} b_j y_{t-j} \right)^2 .$$

La PACF empirique au retard h est le dernier coefficient de cette régression :

$$\widehat{\alpha}(h) = \widehat{a}_h.$$

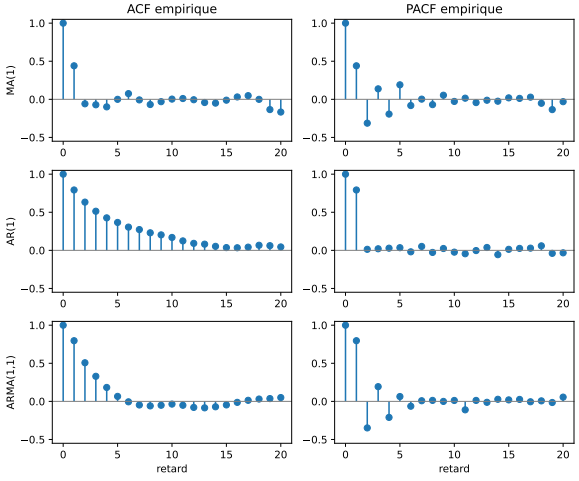
La fonction `pacf()` effectue automatiquement ce calcul pour plusieurs retards.

Signatures classiques

Modèle	**ACF théorique**	**PACF théorique**
MA(q)	coupée après q	décroissance progressive
AR(p)	décroissance progressive	coupée après p
ARMA(p, q)	décroissance progressive	décroissance progressive

Ces signatures donnent une orientation, pas une preuve. Sur une série finie, les graphiques sont bruités.

ACF et PACF : lecture graphique



Lecture qualitative

Une démarche raisonnable :

1. Regarder le chronogramme : la série ressemble-t-elle à une série stationnaire ?
2. Regarder l'ACF : coupure nette ou décroissance progressive ?
3. Regarder la PACF : coupure nette ou décroissance progressive ?
4. Proposer quelques modèles candidats, puis les estimer et les diagnostiquer.

On ne choisit pas définitivement un modèle avec un seul graphique. Les ACF/PACF servent à éviter de partir complètement au hasard.

Sixième partie

Premières commandes R

La simulation permet de relier les formules aux graphiques observés.

Simuler un MA, un AR ou un ARMA

La fonction de base est `arima.sim()`.

```
set.seed(2026)
x.ma   <- arima.sim(model = list(ma = 0.8), n = 240)
x.ar   <- arima.sim(model = list(ar = 0.8), n = 240)
x.arma <- arima.sim(model = list(ar = 0.6, ma = 0.6), n = 240)
```

Dans `model = list(ar = ..., ma = ...)`, on donne les coefficients AR et MA du modèle à simuler.

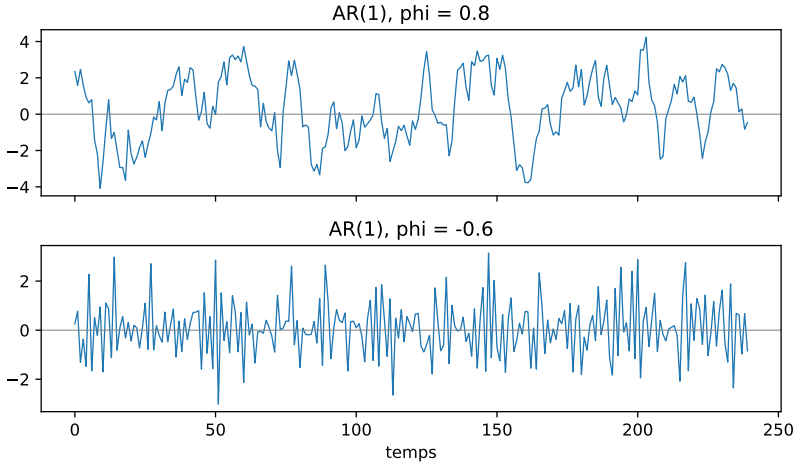
Tracer ACF et PACF empiriques

```
acf(x.ar, lag.max = 20, main = "ACF : AR(1)")
pacf(x.ar, lag.max = 20, main = "PACF : AR(1)")
```

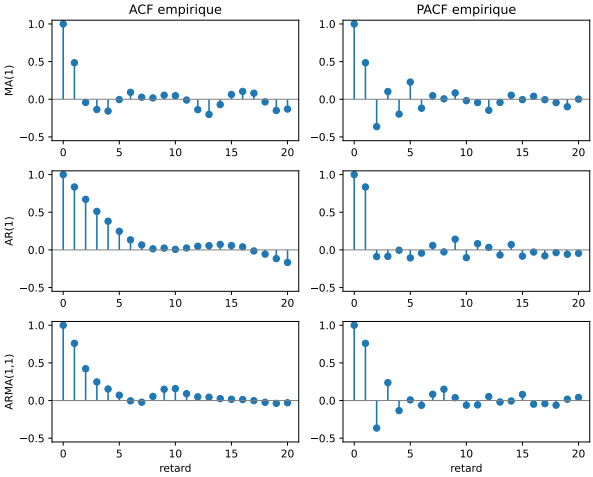
`acf()` trace l'autocorrélation empirique jusqu'au retard demandé.

`pacf()` trace l'autocorrélation partielle empirique.

Deux AR(1) simulés



ACF/PACF empiriques simulées



Synthèse

1. Un MA(q) est une combinaison finie de chocs récents : son ACF théorique est coupée après q.
2. Un AR(p) est une équation récursive : les chocs se propagent dans le temps.
3. Un AR(1) causal existe si $|\phi| < 1$, et son ACF vaut $\rho(h) = \phi^{|h|}$.
4. Pour un AR(p), la causalité se lit sur les racines de $\Phi(z)$.
5. Un ARMA(p, q) combine une partie AR et une partie MA.
6. L'ACF mesure la dépendance totale ; la PACF mesure la dépendance directe après retrait des retards intermédiaires.
7. Les graphiques ACF/PACF donnent des modèles candidats, pas une décision automatique.