

# Calcul des prévisions

Calcul de la meilleure  
prévision linéaire lorsque le  
processus est présumé  
stationnaire

# Objectif principal

- ▶ Dans un premier temps on se situe dans le contexte d'un processus stationnaire  $\{X_t\}$ .
- ▶ On cherche à prédire  $X_{n+m}$ , pour  $m = 1, 2, 3, \dots$ , en utilisant une série chronologique  $X_1, \dots, X_n$ .
- ▶ On parle de prévision, et non pas de ~~prédiction~~.
- ▶ La valeur  $m$  est appelée l'horizon.
- ▶ Dans un premier temps on cherche qu'à faire des prévisions sans formuler un modèle et sans estimer des paramètres.
- ▶ Les prévisions seront typiquement fonction de l'autocorrélation théorique.

# Recherche de la meilleure prévision

► Considérons le critère:

- $E[\{X_{n+m} - g(\mathbf{X})\}^2]$ , où  $\mathbf{X} = (X_1, \dots, X_n)^\top$ , et  $g(\cdot)$  une fonction.
- La meilleure prévision selon le critère précédent est l'espérance conditionnelle  $E(X_{n+m} | \mathbf{X})$ .
- Trouver la meilleure prévision peut être difficile à calculer car il nécessite la connaissance des lois conditionnelles.
- On rappelle qu'un cas notable où l'espérance conditionnelle peut être calculée est dans le cas du processus Gaussien car les lois conditionnelles sont gaussiennes.

# Meilleure prévision versus meilleure prévision linéaire

- ▶ Afin de déterminer  $E(X_{n+m} | \mathbf{X})$ , on a besoin de la loi de  $X_{n+m} | \mathbf{X} = \mathbf{x}$ , qui peut être difficile à déterminer (et ensuite on doit calculer l'espérance de cette loi conditionnelle).
- ▶ Une façon de contourner les difficultés est de se restreindre à la meilleure prévision linéaire:
  - ▶  $P(X_{n+m} | X_1, \dots, X_n) \equiv P_n(X_{n+m}) = \alpha_0 + \sum_{k=1}^n \alpha_k X_k$ .
- ▶ Il reste à déterminer les constantes  $\alpha_0, \alpha_1, \dots, \alpha_n$ , qui sont des nombres réels. On peut adopter le critère de l'erreur quadratique moyenne de prévision:

$$\min_{\alpha_0, \alpha_1, \dots, \alpha_n} [\{X_{n+m} - \alpha_0 - \sum_{k=1}^n \alpha_k X_k\}^2]$$

# Propriétés de la meilleure prévision linéaire

- ▶ La meilleure prévision linéaire n'est fonction que des moments du premier et second ordre (variance et covariances).
- ▶ Bien sûr, la meilleure prévision linéaire est fonction de moments théoriques; cependant, en pratique, on estime ces quantités par les versions empiriques.
- ▶ Comme déjà vu, lorsque le processus  $\{X_t\}$  est Gaussien, la meilleure prévision linéaire est également la meilleure prévision.
  - ▶ Ceci devrait être relativement aisé à voir: on a donné la loi conditionnelle, qui est normale. On calcule alors l'espérance conditionnelle, et l'on constate que cette prévision est en fait linéaire.
- ▶ Un point reste à aborder, à savoir comment déterminer les constantes  $\alpha_0, \alpha_1, \dots, \alpha_k$ . C'est l'objet de la propriété  $P_3$ .

# Propriété P<sub>3</sub>: Calcul de la meilleure prévision linéaire pour $\{X_t\}$ stationnaire

- ▶ Étant donné la série chronologique  $X_t, t = 1, \dots, n$ , la meilleure prévision linéaire de  $X_{n+m}$ , pour un horizon  $m \geq 1$  est de la forme:
  - ▶  $P(X_{n+m} | X_1, \dots, X_n) \equiv P_n(X_{n+m}) = \alpha_0 + \sum_{k=1}^n \alpha_k X_k$ .
- ▶ Les coefficients sont obtenus en résolvant les équations de prévision suivantes:
  - ▶  $E[\{X_{n+m} - P_n(X_{n+m})\}X_k] = 0$ , pour  $k = 0, 1, \dots, n$ .
- ▶ Dans le système précédent, on pose  $X_0 = 1$ .

# Propriétés principales de l'opérateur de prévision $P(\cdot | \mathbf{W})$

- ▶ Posons  $\mathbf{W} = (W_n, \dots, W_1)^\top$  un vecteur de variables aléatoires. Supposons que  $U$  et  $V$  sont des variables aléatoires satisfaisants  $E(U^2) < \infty$  et  $E(V^2) < \infty$ . Dénotons  $P(U | \mathbf{W})$  la meilleure prévision linéaire de  $U$  en fonction de 1 et de  $\mathbf{W}$ . Posons  $\Gamma = \text{var}(\mathbf{W})$  et soient  $\beta, \alpha_1, \dots, \alpha_n$  des constantes.
- ▶ Propriétés fondamentales de l'opérateur de prévision:
- ▶ 1.  $P(U | \mathbf{W}) = E(U) + \mathbf{a}^\top \{\mathbf{W} - E(\mathbf{W})\}$ , où  $\mathbf{a}$  satisfait:  $\Gamma \mathbf{a} = \text{cov}(U, \mathbf{W})$ .
- ▶ 2.  $E[\{U - P(U | \mathbf{W})\} \mathbf{W}] = \mathbf{0}$  et de plus  $E\{U - P(U | \mathbf{W})\} = 0$ .
- ▶ 3.  $E[\{U - P(U | \mathbf{W})\}^2] = \text{var}(U) - \mathbf{a}^\top \text{cov}(U, \mathbf{W})$ .
- ▶ 4.  $P(\alpha_1 U + \alpha_2 V + \beta | \mathbf{W}) = \alpha_1 P(U | \mathbf{W}) + \alpha_2 P(V | \mathbf{W}) + \beta$ .
- ▶ 5.  $P(\sum_{i=1}^n \alpha_i W_i + \beta | \mathbf{W}) = \sum_{i=1}^n \alpha_i W_i + \beta$ .
- ▶ 6.  $P(U | \mathbf{W}) = E(U)$  si  $\text{cov}(U, \mathbf{W}) = \mathbf{0}$ .

# Cas d'un processus stationnaire dont la moyenne est nulle

- ▶ Dans la Propriété P<sub>3</sub>, si l'on pose  $k = 0$ , on obtient l'équation suivante:

- ▶  $E(X_{n+m} - \alpha_0 - \sum_{k=1}^n \alpha_k X_k) = 0.$

- ▶ Autrement dit, si le processus admet  $E(X_k) = \mu$ , cette équation peut être reformulée comme:

- ▶  $\mu - \alpha_0 - \sum_{k=1}^n \alpha_k \mu = 0.$

- ▶ Ainsi,  $\alpha_0 = \mu(1 - \sum_{k=1}^n \alpha_k).$

- ▶ Donc la meilleure prévision linéaire dans le cas de  $E(X_k) = \mu$  s'exprime comme:

- ▶  $P_n(X_{n+m}) = \mu + \sum_{k=1}^n \alpha_k (X_k - \mu).$

- ▶ Si  $\mu = 0$ , alors  $\alpha_0 = 0$  et  $P_n(X_{n+m}) = \sum_{k=1}^n \alpha_k X_k.$



# Cas particulier important: prévision d'horizon un

- ▶ On note que la propriété  $P_3$  est valable pour un horizon arbitraire.
- ▶ Un cas est particulièrement important: le cas  $m = 1$ .
- ▶ Afin de simplifier, on suppose que  $\mu = 0$ .
- ▶ Il est d'usage de noter la prévision d'horizon un comme suit:
  - ▶  $P_n(X_{n+1}) = \phi_{n1}X_n + \dots + \phi_{nn}X_1$ .
- ▶ Une attention particulière sera portée sur les coefficients  $\phi_{nn}$ ,  $n \geq 1$ .

# Cas particulier important: prévision d'horizon un (suite)

- ▶ Rappel:  $P_n(X_{n+1}) = \phi_{n1}X_n + \dots + \phi_{nn}X_1$ .
- ▶ Selon la propriété P<sub>3</sub>, il faut résoudre le système suivant:
  - ▶  $E\left(\left\{X_{n+1} - \sum_{k=1}^n \phi_{nk}X_{n-k+1}\right\}X_{n+1-k}\right) = 0$ , pour  $k = 1, \dots, n$ .
- ▶ Puisque  $\{X_t\}$  est stationnaire, le système s'écrit comme le système d'équations:
  - ▶  $\gamma(k) - \sum_{j=1}^n \phi_{nj}\gamma(k-j) = 0$ ,  $k = 1, \dots, n$ .
- ▶ Une solution itérative permet de trouver les coefficients de la meilleure prévision linéaire: c'est l'algorithme de Durbin-Levinson.

# Propriété P<sub>4</sub>: Algorithme de Durbin-Levinson

- ▶ Soit  $\{X_t\}$  un processus stationnaire de moyenne zéro. Posons  $P_n(X_{n+1}) = \phi_{n1}X_n + \dots + \phi_{nn}X_1$ . On pose également l'erreur quadratique moyenne de prévision:

$$\text{▶ } v_n = E[\{X_{n+1} - P_n(X_{n+1})\}^2]$$

- ▶ Les coefficients  $\phi_{nj}$  et les erreurs quadratiques moyennes de prévision peuvent être calculés récursivement comme suit:

$$\text{▶ } \phi_{00} = 1, \text{ et } v_1^{(0)} = \gamma(0).$$

- ▶ Pour  $n \geq 1$ :

$$\text{▶ } \phi_{nn} = \frac{\gamma(n) - \sum_{k=1}^{n-1} \phi_{n-1,k} \gamma(n-k)}{\gamma(0) - \sum_{k=1}^{n-1} \phi_{n-1,k} \gamma(k)}$$

- ▶ L'erreur quadratique moyenne de prévision est  $v_{n+1}^{(n)} = v_n^{(n-1)}(1 - \phi_{nn}^2)$ .

- ▶ Pour  $n \geq 2$ :

$$\text{▶ } \phi_{nk} = \phi_{n-1,k} - \phi_{nn} \phi_{n-1,n-k}, \quad k = 1, \dots, n-1.$$

# Exemple: Mise en œuvre de l'algorithme de Durbin-Levinson

► On initialise avec  $\phi_{00} = 1$ , et  $v_1 = \gamma(0)$ .

► Pour  $n = 1$ :

$$\text{► } \phi_{11} = \frac{\gamma(1) - 0}{\gamma(0) - 0} = \rho(1).$$

$$\text{► } v_2 = v_1(1 - \phi_{11}^2) = \gamma(0)\{1 - \rho^2(1)\}.$$

► Pour  $n = 2$ :

$$\text{► } \phi_{22} = \frac{\gamma(2) - \gamma(1)\phi_{11}}{\gamma(0) - \phi_{11}\gamma(1)} = \frac{\gamma(2) - \gamma(1)\rho(1)}{\gamma(0)\{1 - \rho^2(1)\}} = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)} = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}}.$$

$$\text{► } \phi_{21} = \phi_{11} - \phi_{11}\phi_{22} = \phi_{11}(1 - \phi_{22}) = \rho(1)(1 - \phi_{22}).$$

$$\text{► } v_3 = v_2(1 - \phi_{22}^2) = \gamma(0)(1 - \phi_{11}^2)(1 - \phi_{22}^2)$$

► Etc.

## Propriété P<sub>5</sub>: Solution itérative pour la fonction d'autocorrélation partielle

- ▶ L'algorithme de Durbin-Levinson (Propriété P<sub>4</sub>) a pour intérêt principal la Propriété P<sub>5</sub>, qui s'énonce comme suit:
- ▶ Considérons un processus stationnaire  $\{X_t\}$ . Alors les coefficients  $\phi_{nn}$ , pour  $n \geq 1$ , dans la Propriété P<sub>4</sub> sont les autocorrélations partielles.

# Prévision d'horizon $m$

- ▶ La meilleure prévision d'horizon  $m$  est notée (car  $\mu = 0$ ):

- ▶  $P_n(X_{n+m}) = \phi_{n1}^{(m)} X_n + \dots + \phi_{nn}^{(m)} X_1.$

- ▶ Selon la Propriété P<sub>3</sub>, le système à résoudre est de la forme:

- ▶  $E \left( \left\{ X_{n+m} - \sum_{k=1}^n \phi_{nk}^{(m)} X_{n-k+1} \right\} X_{n+1-k} \right) = 0, \text{ pour } k = 1, \dots, n.$

- ▶ Puisque  $\{X_t\}$  est stationnaire, le système s'écrit comme le système d'équations:

- ▶  $\gamma(m+k-1) - \sum_{j=1}^n \phi_{nj}^{(m)} \gamma(k-j) = 0, k = 1, \dots, n.$

- ▶ En utilisant des arguments analogues au cas  $m = 1$ , on montre que l'erreur quadratique moyenne de prévision s'écrit:

- ▶  $v_n = E[\{X_{n+m} - P_n(X_{n+m})\}^2] = \gamma(0) - \gamma_n^{m\top} \Gamma_n^{-1} \gamma_n^m,$

- ▶  $\gamma_n^m = \begin{pmatrix} \gamma(m) \\ \vdots \\ \gamma(m+n-1) \end{pmatrix}; \Gamma_n = \begin{pmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{pmatrix}.$

# Algorithme des innovations

- ▶ Les propriétés  $P_4$  et  $P_5$  font ressortir dans une certaine mesure les propriétés des processus autorégressifs, du fait que les autocorrélations partielles finissent par s'annuler.
- ▶ Un autre algorithme, l'algorithme des innovations, présente certains avantages dans le calcul des prévisions des moyennes-mobiles, donc dans le cadre des modèles  $MA(q)$ .
- ▶ Il est valide en général, pour un processus  $\{X_t\}$  tel que  $E(X_t^2) < \infty$ . Ainsi, dans sa version la plus générale, le processus pourrait ne pas être stationnaire. La propriété  $P_6$  dans l'ouvrage de référence se limite au cas stationnaire. On va présenter le cas général.

# Propriété P<sub>6</sub>: Algorithme des innovations

- ▶ Pour un processus de moyenne zéro, définissons  $\kappa(i, j) = E(X_i X_j)$ .
- ▶ Il est utile d'introduire:  $\hat{X}_n = \begin{cases} 0, & n = 1, \\ P_{n-1}(X_n), & n > 1. \end{cases}$
- ▶ Comme précédemment,  $v_n = E[\{X_{n+1} - P_n(X_{n+1})\}^2]$ .
- ▶ L'algorithme des innovations écrit la meilleure prévision linéaire de la manière suivante:
  - ▶  $\hat{X}_{n+1} = \begin{cases} 0, & n = 0, \\ \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}), & n > 0. \end{cases}$
- ▶ Les coefficients  $\theta_{nj}, \dots, \theta_{nn}$  peuvent être calculés récursivement.



# Propriété P<sub>6</sub>: Algorithme des innovations (suite)

- ▶ Rappel:  $\hat{X}_{n+1} = \begin{cases} 0, & n = 0, \\ \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}), & n > 0. \end{cases}$
- ▶ Les coefficients  $\theta_{nj}, \dots, \theta_{nn}$  peuvent être calculés récursivement de la manière suivante:
  - ▶  $v_0 = \kappa(1,1)$ ,
  - ▶  $\theta_{n,n-k} = v_k^{-1} \{ \kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \}$ , pour  $0 \leq k < n$ .
  - ▶ De plus,  $v_n = \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j$
- ▶ L'algorithme fonctionne comme suit: On calcule  $v_0$ ; ensuite  $\theta_{11}$  et  $v_1$ ; ensuite  $\theta_{22}, \theta_{21}$  et  $v_2$ ; ensuite  $\theta_{33}, \theta_{32}, \theta_{31}$  et  $v_3$ . Etc.