CORRÉLOGRAMME (C5, D2, F3, N)
(02 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **corrélogramme** est un instrument probabiliste qui permet l'étude, dans l'**espace des états**, des corrélations internes à un processus ou à une série temporelle. Son utilisation se limite, le plus souvent, aux processus ou séries stationnaires en covariance (cf **processus stationnaire en covariance**). On distingue entre corrélogramme théorique (celui du processus) et corrélogramme empirique (celui de la série).

(i) Soit $X = (X_t)_{t \in T}$ un **processus** réel scalaire tq $T \subset \mathbf{R}$: eg $T = N_T^*$, ou $T = \mathbf{N}$, $T = \mathbf{Z}$, $T = \mathbf{Q}$, $T = \mathcal{I}(\mathbf{R})$ (intervalles de $\mathbf{R}$) ou $T = \mathbf{R}$. On suppose X stationnaire en covariance et l'on pose :

$$\begin{aligned} &E\, X_t = \mu, && \forall\, t \in T, \\ (1) \quad &V\, X_t = E\,(X_t - \mu)^2 = \sigma^2, && \forall\, t \in T, \\ &C\,(X_s, X_t) = E\,(X_s - \mu)(X_t - \mu) = \gamma_{st}, && \forall\, (s, t) \in T^2. \end{aligned}$$

La **fonction d'autocovariance** (théorique) de X étant définie par :

$$(2) \qquad \gamma : t - s \mapsto \gamma_{st}, \qquad \forall\, (s, t) \in T^2,$$

la **fonction d'autocorrélation** (théorique) de X s'en déduit selon :

$$(3) \qquad \rho : t - s \mapsto \rho_{st} = \gamma_{st} / (\gamma_{ss} \,.\, \gamma_{tt})^{1/2}, \quad \forall\, (s, t) \in T^2.$$

On appelle **corrélogramme** (théorique) de X le **graphe** de la fonction d'autocorrélation de X.

(ii) La fonction d'autocovariance γ est à valeurs dans $\mathbf{R}$, et la fonction d'autocorrélation ρ dans [-1, +1]. Comme X est stationnaire au second ordre, les **moments** du second ordre ne dépendent que de la différence $t - s = \theta$. C'est pourquoi l'on note souvent (en particulier lorsque $T = \mathbf{Z}$) :

$$(4) \qquad \begin{aligned} &\gamma_\theta = \gamma\,(\theta) = \gamma_{s,t-\theta}, && \forall\, (s, t-\theta) \in T^2, \\ &\rho_\theta = \rho\,(\theta) = \gamma\,(\theta) / \gamma\,(0) = \gamma_{s,t-\theta} / \gamma_{ss}, && \forall\, (s, t-\theta) \in T^2. \end{aligned}$$

Le corrélogramme correspond ainsi à une fonction d'autocorrélation paire ($\rho_{-\theta} = \rho_\theta$). Sa représentation graphique peut alors se restreindre aux seuls indices $\theta \in \mathbf{N}$. On a aussi $|\rho_\theta| \leq 1$, $\forall\, \theta \in \mathbf{N}$.

(iii) Soit $x = (x_t)_{t=1,\ldots,T}$ une **série temporelle**. L'**autocovariance (empirique)** étant définie selon :

$$(5) \qquad c_\theta = (T - \theta)^{-1} \Sigma_{t=1}^{T-\theta} (x_t - \overline{x}_T)(x_{t+\theta} - \overline{x}_T), \qquad \forall\, \theta \in N_{T-1},$$

où $\bar{x}_T = T^{-1} \Sigma_{t=1}^{T} x_t$ (**moyenne empirique**), l'**autocorrélation (empirique)** s'en déduit selon :

(6) $\quad r_\theta = c_\theta / c_0 , \qquad \forall \theta \in N_{T-1}$.

Le graphe de la fonction $\theta \mapsto r_\theta$ est appelé **corrélogramme (empirique)** de x (ou de X) et l'on pose $c_{-\theta} = c_\theta$ et $r_{-\theta} = r_\theta$.

(iv) Le **processus générateur de données**, ie celui qui génère x, étant stationnaire en covariance, les statistiques $\bar{x}_T$, c_θ et r_θ sont souvent utilisées comme **estimateurs** resp des **paramètres** μ, γ_θ et ρ_θ (cf **statistique naturelle**).

Il en existe des formes modifiées (cf **modification**), dont :

$$(7) \quad \begin{aligned} c_\theta' &= (T - \theta)^{-1} . \{\Sigma_{t=1}^{T-\theta} x_t x_{t+\theta} - (T - \theta)^{-1} . (\Sigma_{t=\theta+1}^{T} x_t) . (\Sigma_{t=1}^{T-\theta} x_t)\}, \\ r'_\theta &= c_\theta' / c_0', \end{aligned}$$

et le graphe $\theta \mapsto r_\theta'$ sert aussi d'estimateur du corrélogramme.

Les estimateurs précédents sont des **estimateurs convergents** (en probabilité) pour certaines classes de processus, notamment les processus ergodiques : eg le **processus linéaire** (cf **ergodicité**, **théorie ergodique**).

(v) Lorsque X est un processus scalaire complexe, stationnaire en covariance, on peut définir :

$$(8) \quad \begin{aligned} E\, X_t &= \mu \in \mathbf{C}, \\ C\, (X_s , \bar{X}_t) &= E\, (X_s - \mu)(\bar{X}_t - \bar{\mu}) = \gamma_{st} \in \mathbf{C}, \end{aligned}$$

où $\bar{z}$ est le nombre complexe conjugué de z. Cette représentation assure la réalité de $V\, X_t = \gamma_{tt}$. On en déduit la définition du **corrélogramme (à valeurs complexes)** : $t - s \mapsto \rho_{st} = \gamma_{st} / \gamma_{ss}$.

(vi) De même, lorsque X est vectoriel (réel ou complexe), on peut définir :

$$(9) \quad \begin{aligned} E\, X_t &= \mu \in \mathbf{C}^K, \\ \gamma_{st} &= C\, (X_s , \bar{X}_t) = E\, (X_s - \mu)(\bar{X}_t - \bar{\mu})' \in M_K\, (\mathbf{C}), \end{aligned}$$

ce qui assure la réalité de $V\, X_t = \gamma_{tt}$.

Mais la notion de corrélogramme ne s'étend pas aisément à ce cas. On ne peut définir que des corrélogrammes relatifs à chaque coordonnée de X (cf **corrélogramme croisé**).