TP Séries Chronologiques n^0 6 : Processus ARMA

L'objectif de ce TP est d'estimer les paramètres d'un processus ARMA, d'estimer son ordre et enfin de faire de la prédiction.

Estimation des paramètres et de l'ordre d'un ARMA

Splus -e	Lancer Splus (avec l'option éditeur de ligne)
help.start()	Ouvre la fenêtre d'aide.
motif()	Ouverture d'une fenêtre graphique.
	Corrélogrammes
	Simuler une trajectoire de taille 100, $X100$, du processus ARMA $X=(X_n)_{n\in\mathbb{Z}}$ tel que : $(1-0.4B).(1-0.5B)X=\varepsilon$, où B est l'opérateur retard et $\varepsilon=(\varepsilon)_{n\in\mathbb{Z}}$ est un bruit blanc gaussien de variance 2.
$acf(X100)$	Corrélogramme de $X100$. Déterminer le corrélogramme théorique de $X100$. Comparer.
$acf(X100, type = '' p''))$	Fonction de corrélation partielle empirique de $X100$. Déterminer théoriquement les valeurs et comparer. A votre avis, pour identifier X, quel "corrélogramme" vaut-il mieux utiliser? Augmenter la taille de la trajectoire et recommencer. Conclusions ?
	Estimation des paramètres et de l'ordre par minimum de contraste
	Estimer les paramètres de $X100$ pour une trajectoire de taille 100, à partir d'une estimation par minimum de contraste de Whittle (taper $param\_arima.mle(...)$) en précisant un ordre $(2,0,0)$ (le 0 du milieu correspond à un paramètre distingant les ARIMA, que l'on prendra par la suite toujours égal à 0). Vérifier la convergence et donner la matrice des estimateurs des paramètres.
$arima.diag(param)$	Reconstruit le bruit blanc à partir des coefficients estimés, c'est-à-dire que si $P(B).X=Q(B).\varepsilon$, l'estimation par MLE permet d'obtenir $\widehat{P}$ et $\widehat{Q}$, et donc $\dfrac{\widehat{P}(B)}{\widehat{Q}(B)}X100=\widehat{\varepsilon}100$. On vérifie alors si le bruit blanc "reconstitué" $\widehat{\varepsilon}100$ est bien un bruit blanc. La convergence de la procédure $arima.mle$ et la "blancheur" de $\widehat{\varepsilon}100$ permettent de vérifier si l'ordre choisi était le bon. Recommencer ces procédures en choisissant pour ordre $(1,0,0)$, $(0,0,1)$, $(0,0,2)$, $(2,0,1)$, $(3,0,3)$. Conclusions ? Recommencer en passant à la taille 1000.
	Estimation des paramètres et de l'ordre à partir des corrélations
	On travaille à partir de $X100$ déjà simulé et on estime les paramétresb et l'ordre.
$MA\_acf(X100)$	À partir du corrélogramme, on peut estimer l'ordre d'un MA modélisant $X100$ par l'entier k tel que les coefficients de corrélation $\rho(j)$,$j=0,1,..,k$ sont considérés comme non nuls.
$AR\_acf(X100, type = '' p'')$	À partir du corrélogramme partiel, on peut estimer l'ordre d'un AR modélisant $X100$ par l'entier k tel que les k premiers coefficients de corrélation partielle sont considérés comme non nuls. Conclusions ? Recommencer avec la taille 1000.
$ParamAR\_ar(X100)$	Cette fois-ci, on fait comme si la série $X100$ était un AR et on estime les coefficients de cet AR à partir des équations de Yule-Walker (utilisant les corrélations empiriques). Que trouvez-vous? (ordre, coefficients)? Considérer la série des résidus ($ParamAR\$res$), c'est-à-dire le bruit blanc reconstitué à partir de $X100$ et des coefficients estimés. Est-ce un bruit blanc? Quelle estimation des paramètres donne les meilleurs résultats (plusieurs trajectoires...) ?

Exercice

Simuler une trajectoire de taille 500 du processus $Y = (Y_n)_{n\in\mathbb{Z}}$ tel que $Y_{n+1} = -0.8Y_n + \varepsilon_n - 0.9\varepsilon_{n-1}$, avec $\varepsilon = (\varepsilon)_{n\in\mathbb{Z}}$ est un bruit blanc gaussien de variance 1. Quel type de processus est Y ? Estimer l'ordre de ce processus à partir des procédures indiquées précédemment. Conclusions ?

Prédiction pour des processus ARMA

	Prédictions pour un processus ARMA
	Considérer Z comportant les 490 premières valeurs de Y précédemment simulée. Estimer les coefficients par maximum de vraisemblance approchée (taper $ParaZ\_arima.mle(...)$), avec l'ordre $(1,0,1)$.
	Considérer alors $ParaZ\$model$. Quelle est cette variable?
	On veut prévoir les 10 données suivantes de Z à partir du modèle estimé précédent.
pZ_arima.forecast(Z,ParaZ$model,10)	On a ainsi estimé les 10 dernières données de Y à partir des 490 premières.
pZ$mean	Valeurs prévues pour des 10 données.
pZ$std	Ecarts-type des estimateurs de ces 10 données.
	Quelle est la loi de $\widehat{Y}_{491}$?
	Comparer les 10 données prévues avec les 10 valeurs réelles. Conclusion?
	Expliquer la convergence des valeurs prédites.
	Faire la même chose en modélisant Y par un AR (remplacer la commande $ParaY\$model$ par $list(ar = arY\$ar)$). A-t-on améliorer la prédiction?