ECN 6578A, Économétrie des marchés financiers, Hiver 2023

Cours 4

William McCausland

2023-01-30

Plan

1. Prévision linéaire avec un modèle ARMA(p,q)
2. Modèles pour les deux premiers moments conditionnels (GARCH)
3. Simulation des modèles GARCH
4. Vraisemblance des modèles GARCH

Prévison avec un modèle ARMA(p,q)

- Le problème : prévoir r_{t+h} à t, mesurer l'incertitude.
- Types de prévision
  - ponctuelle,
  - par intervalle,
  - par densité
- Si l'objectif est de choisir $\hat{r}_t(h)$ pour minimiser $E[(r_{t+h} - \hat{r}_t(h))^2 | F_t]$, la meilleure prévision ponctuelle est $\hat{r}_t(h) = E[r_{t+h}|F_t]$ et la valeur minimal est $\mathrm{Var}[r_{t+h}|F_t]$.
- Dans un modèle ARMA(p,q), $E[r_{t+h}|F_t]$ est linéaire en $r_t, r_{t-1}, \ldots, r_{t-p}$ et $a_t, a_{t-1}, \ldots, a_{t-q}$.
- Avec les coefficients ϕ_i et θ_i, on peut évaluer $E[r_{t+h}|F_t]$.
- Avec σ_a^2 aussi on peut évaluer $\mathrm{Var}[r_{t+h}|F_t]$.
- Pour simplifier un peu, F_t comprend $r_t, r_{t-1}, \ldots$ et $a_t, a_{t-1}, \ldots$.
- En réalité, on observe un échantillon $r_1, \ldots, r_T$ et recouvre $a_1, \ldots, a_T$ seulement de façon approximative.

Prévision avec un modèle ARMA(2,2) à un horizon $h = 1$

- Équation ARMA(2,2) pour r_{t+1} :

$$r_{t+1} = \phi_1 r_t + \phi_2 r_{t-1} + a_{t+1} - \theta_1 a_t - \theta_2 a_{t-1}.$$

- Prenez l'espérance conditionnelle $E[\cdot|F_t]$ des deux côtés pour obtenir une prévision :

$$E[r_{t+1}|F_t] = \phi_1 r_t + \phi_2 r_{t-1} - \theta_1 a_t - \theta_2 a_{t-1}.$$

- Calculez l'erreur de la prévision :

$$r_{t+1} - E[r_{t+1}|F_t] = a_{t+1}.$$

- Calculez la variance de l'erreur de la prévision :

$$\mathrm{Var}[r_{t+1}|F_t] = \mathrm{Var}[a_{t+1}|F_t] = \sigma_a^2.$$

Prévision à un horizon $h = 2$: la prévision ponctuelle

- Équation pour r_{t+2} :

$$r_{t+2} = \phi_1 r_{t+1} + \phi_2 r_t + a_{t+2} - \theta_1 a_{t+1} - \theta_2 a_t.$$

- Prenez l'espérance conditionnelle $E[\cdot|F_t]$ des deux côtés pour obtenir :

$$E[r_{t+2}|F_t] = \phi_1 E[r_{t+1}|F_t] + \phi_2 r_t - \theta_2 a_t,$$

où $E[r_{t+1}|F_t] = \phi_1 r_t + \phi_2 r_{t-1} - \theta_1 a_t - \theta_2 a_{t-1}$.

- Alors la prévision à $h = 2$ est

$$\begin{aligned} E[r_{t+2}|F_t] &= \phi_1[\phi_1 r_t + \phi_2 r_{t-1} - \theta_1 a_t - \theta_2 a_{t-1}] + \phi_2 r_t - \theta_2 a_t \\ &= (\phi_1^2 + \phi_2) r_t + \phi_1 \phi_2 r_{t-1} - (\phi_1 \theta_1 + \theta_2) a_t - \phi_1 \theta_2 a_{t-1}. \end{aligned}$$

Variance de l'erreur ($h = 2$) : 1er perspectif de trois

- Calculez l'erreur de la prévision :

$$\begin{aligned} r_{t+2} - E[r_{t+2}|F_t] &= \phi_1(r_{t+1} - E[r_{t+1}|F_t]) + a_{t+2} - \theta_1 a_{t+1} \\ &= (\phi_1 - \theta_1)a_{t+1} + a_{t+2}. \end{aligned}$$

- Calculez la variance de l'erreur de la prévision :

$$\mathrm{Var}[r_{t+2}|F_t] = [(\phi_1 - \theta_1)^2 + 1]\sigma_a^2.$$

Variance de l'erreur ($h = 2$) : 2e perspectif de trois

- Par la loi de variance totale :

$$\mathrm{Var}[r_{t+2}|F_t] = E[\mathrm{Var}[r_{t+2}|F_{t+1}]|F_t] + \mathrm{Var}[E[r_{t+2}|F_{t+1}]|F_t]$$

- Puisque $\mathrm{Var}[r_{t+2}|F_{t+1}] = \sigma_a^2$,

$$E[\mathrm{Var}[r_{t+2}|F_{t+1}]|F_t] = \sigma_a^2.$$

- Puisque $E[r_{t+2}|F_{t+1}] = \phi_1 r_{t+1} + \phi_2 r_t - \theta_1 a_{t+1} - \theta_2 a_t$,

$$\mathrm{Var}[E[r_{t+2}|F_{t+1}]|F_t] = \mathrm{Var}[(\phi_1 - \theta_1)a_{t+1}|F_t] = (\phi_1 - \theta_1)^2\sigma_a^2$$

- Alors,

$$\mathrm{Var}[r_{t+2}|F_t] = [1 + (\phi_1 - \theta_1)^2]\sigma_a^2.$$

Variance de l'erreur ($h = 2$) : 3e perspectif de trois

- La représentation MA infinie donne

$$r_{t+2} = a_{t+2} + \psi_1 a_{t+1} + \psi_2 a_t + \dots$$

- La variance conditionnelle est donc

$$\begin{aligned}
\mathrm{Var}[r_{t+2}|F_t] &= \mathrm{Var}[a_{t+2} + \psi_1 a_{t+1}|F_t] \\
&= (1 + \psi_1^2)\sigma_a^2 \\
&= [1 + (\phi_1 - \theta_1)^2]\sigma_a^2.
\end{aligned}$$

Le modèle ARMA(p,q) comme modèle pour la moyenne conditionnelle

- L'équation ARMA(p,q) :

$$r_t = \phi_1 r_{t-1} + \ldots + \phi_p r_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}.$$

- La moyenne conditionnelle :

$$\mu_t \equiv E[r_t|F_{t-1}] = \phi_1 r_{t-1} + \ldots + \phi_p r_{t-1} - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}.$$

- Notez que l'innovation est une erreur de prévision :

$$a_t = r_t - E[r_t|F_{t-1}].$$

- Si on connait $r_{t-1}, \ldots, r_{t-p}$ et $a_{t-1}, \ldots, a_{t-q}$, on apprend a_t en même temps qu'on observe r_t.

Introduction aux modèles ARCH, GARCH

- Modèles pour les deux premiers moments conditionnels de r_t.
- F_{t-1} représente toute l'information connue à $t-1$.
- Au minimum, F_{t-1} comprend $r_{t-1}, r_{t-2}, \ldots$.
- Définitions de la moyenne, de la variance conditionnelle de r_t :

$$\mu_t \equiv E[r_t|F_{t-1}], \qquad \sigma_t^2 \equiv \mathrm{Var}[r_t|F_{t-1}].$$

- 'Hétéroscédasticité conditionnelle' veut dire que σ_t^2 varie avec t.
- Convention alternative (qu'on n'utilise pas ici) où l'indice indique le moment où la quantité est connue : $\mu_{t-1} \equiv E[r_t|F_{t-1}]$.
- Autres définitions : $a_t \equiv r_t - \mu_t$, $\epsilon_t \equiv a_t/\sigma_t$, alors

$$r_t = \mu_t + \sigma_t \epsilon_t.$$

- Notez que $\epsilon_t|F_{t-1} \sim (0, 1)$.

Exemple ARMA(p,q)-GARCH(m,s)

- Un modèle pour μ_t : (cas spécial de 3.3)

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j}.$$

- Un modèle pour σ_t^2 : (3.14)

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2.$$

- $r_t | F_{t-1} \sim (\mu_t, \sigma_t^2)$

Modèle ARCH

- Le modèle ARCH(m) d'Engle :

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2,$$

$$a_t = \sigma_t \epsilon_t, \quad \epsilon_t | F_{t-1} \sim (0, 1).$$

- La spécification implicite de μ_t ici est $\mu_t = 0$.
- Puisque $r_t \equiv \mu_t + a_t$, r_t et a_t ici sont pareille.
- La spécification de σ_t^2 dépend du passé du processus.
- On suppose que a_t est covariance stationnaire, ce qui entraîne des restrictions aux paramètres.
- σ_t connu à $t - 1$; a_t et ϵ_t connus à t.

Moyenne et variance inconditionnelle du modèle ARCH(1)

- ARCH(1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2, \quad E[a_t|F_{t-1}] = 0, \quad \mathrm{Var}[a_t|F_{t-1}] = \sigma_t^2.$$

- Moyenne inconditionnelle

$$E[a_t] = E[E[a_t|F_{t-1}]] = E[0] = 0.$$

- Variance inconditionnelle

$$\mathrm{Var}[a_t] = E[a_t^2] = E[E[a_t^2|F_{t-1}]] = E[\alpha_0 + \alpha_1 a_{t-1}^2],$$

$$(1-\alpha_1)\mathrm{Var}[a_t] = \alpha_0, \qquad \mathrm{Var}[a_t] = \alpha_0/(1-\alpha_1).$$

- Il faut que $\alpha_1 \neq 0$ pour capturer l'hétéroscédasticité conditionnelle, $\alpha_1 > 0$ pour la persistence de la volatilité.
- Il faut que $\alpha_1 < 1$ par covariance stationnarité.
- Il faut que $\alpha_0 > 0$ et $\alpha_1 \geq 0$ par positivité de la variance conditionnelle.

Asymétrie conditionnelle et l'effet de levier

- L'asymétrie conditionnelle est $E[r_t^3 | F_{t-1}]/\sigma_t^3$.
- L'effet de levier est souvent exprimé comme $\mathrm{Cov}[(\sigma_{t+1}^2 - \sigma_t^2), a_t] < 0$.
- Il faut avoir un modèle pour parler des moments conditionnels en général, et les faits empiriques concernant l'asymétrie conditionnelle et l'effet de levier en particulier.

L'asymétrie d'un modèle ARCH(1)

- Souvent les modèles de volatilité spécifie $E[a_t^3|F_{t-1}]$ comme étant zéro ou constant; souvent, ils ne le spécifient pas.
- L'effet de levier en termes du troisième moment inconditionnel :

$$E[(\sigma_{t+1}^2 - \sigma_t^2)a_t] = E[(\alpha_0 + \alpha_1 a_t^2 - \sigma_t^2)a_t] = \alpha_1 E[a_t^3].$$

- Troisième moment inconditionnel en termes du troisième moment conditionnel :

$$E[a_t^3] = E[E[a_t^3|F_{t-1}]].$$

- Si l'asymétrie conditionnelle est nulle, l'asymétrie inconditionnelle et $\mathrm{Cov}[(\sigma_{t+1}^2 - \sigma_t^2), a_t]$ sont nulles aussi.
- Conclusion: pour un ARCH(1), une asymétrie inconditionnelle ou un effet de levier doit passer par une asymétrie conditionnelle. Ces deux effets ne sont pas entrainés par le modèle.
- Même chose pour un ARCH(m).

Aplatissement d'un ARCH(1) gaussien

- Le 4ième moment conditionnel dépend du choix de la loi de a_t.
- Pour un modèle ARCH(1) gaussien,

$$E[a_t^4|F_{t-1}] = 3E[\sigma_t^4|F_{t-1}] = 3(\alpha_0 + \alpha_1 a_{t-1}^2)^2.$$

$$E[a_t^4] = 3E[\alpha_0^2 + 2\alpha_0\alpha_1 a_{t-1}^2 + \alpha_1^2 a_{t-1}^4]$$

- Si $E[a_t^4] = E[a_{t-1}^4]$ (une conséquence de stationnarité, mais pas de covariance-stationnarité) alors

$$E[a_t^4] = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}.$$

- Il faut que $\alpha_1^2 < 1/3$ pour l'existence de l'aplatissement. Un peu inflexible.
- L'aplatissement inconditionnel, s'il existe, est de

$$\frac{E[a_t^4]}{E[a_t^2]^2} = 3(1-\alpha_1^2)/(1-3\alpha_1^2) > 3.$$

Les autocorrélations

- Autocorrélation de première ordre de r_t ou a_t:

$$E[a_t a_{t-1}] = E[E[a_t a_{t-1}|F_{t-1}]] = E[a_{t-1}E[a_t|F_{t-1}]] = E[a_{t-1}0] = 0.$$

- Autocorrélation de a_t^2:

$$\begin{aligned} E[a_t^2 a_{t-1}^2] &= E[E[a_t^2 a_{t-1}^2|F_{t-1}]] = E[a_{t-1}^2 E[a_t^2|F_{t-1}]] \\ &= E[a_{t-1}^2(\alpha_0 + \alpha_1 a_{t-1}^2)] = \frac{\alpha_0^2}{1-\alpha_1} + \alpha_1 E[a_{t-1}^4] \end{aligned}$$

$$\mathrm{Cov}[a_t^2, a_{t-1}^2] = E[a_t^2 a_{t-1}^2] - E[a_t^2]E[a_{t-1}^2] = \alpha_1 E[a_{t-1}^4] - \frac{\alpha_0^2\alpha_1}{(1-\alpha_1)^2}$$

$$\mathrm{Var}[a_t^2] = E[a_t^4] - E[a_t^2]^2 = E[a_t^4] - \frac{\alpha_0^2}{(1-\alpha_1)^2}$$

- Avec stationnarité en 4e moment, $\mathrm{corr}[a_t^2, a_{t-1}^2] = \alpha_1$.

Résumé des conclusions théoriques

Le modèle ARCH

- capture la variabilité et la persistance de la volatilité
- capture l'aplatissement inconditionnel plus grand que 3
- un peu d'inflexibilité pour l'autocorrélation de a_t^2
- ne capture pas la longue mémoire pour la volatilité
- ne capture pas l'asymétrie inconditionnelle, indépendamment de l'asymétrie conditionnelle
- ne capture pas l'effet de levier, indépendamment de l'asymétrie conditionnelle

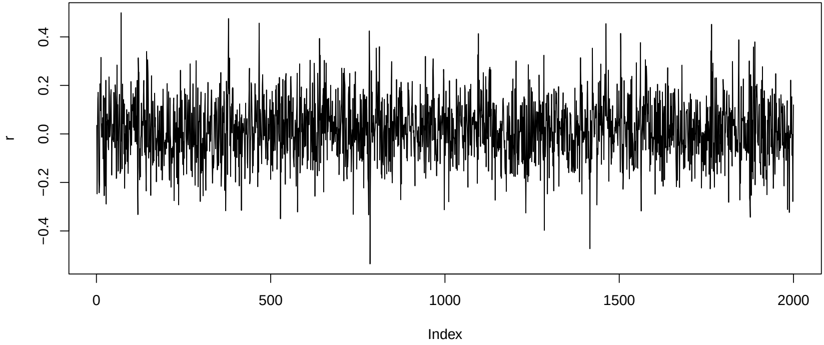
Simulation du modèle ARCH(3)

```r
# Paramètres (de l'exemple Intel, Exemple 3.1)
mu = 0.0122
al0 = 0.0106; al = c(0.2131, 0.0770, 0.0599)
variance = al0/(1-al[1]-al[2]-al[3])

T = 2000 # Nombre de périodes
epsilon = rnorm(T) # Innovations gaussiennes
a = rep(0, T); r = rep(0, T) # Mémoire réservé
a[1:3] = rnorm(3, sd=sqrt(variance));
r[1:3] = a[1:3] + mu
for (t in 4:T) {
  mu_t = mu
  sigma2_t = al0 + al %*% a[(t-1):(t-3)]^2
  a[t] = sqrt(sigma2_t) * epsilon[t]
  r[t] = a[t] + mu_t
}
```

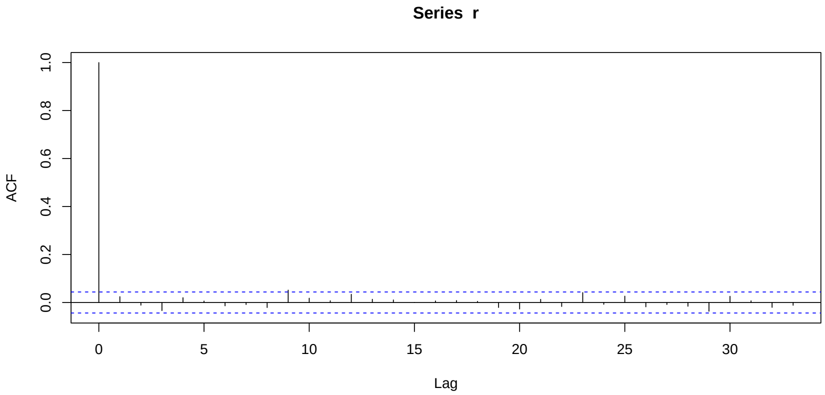
Graphique de r_t artificiel

```
plot(r, type='l')
```



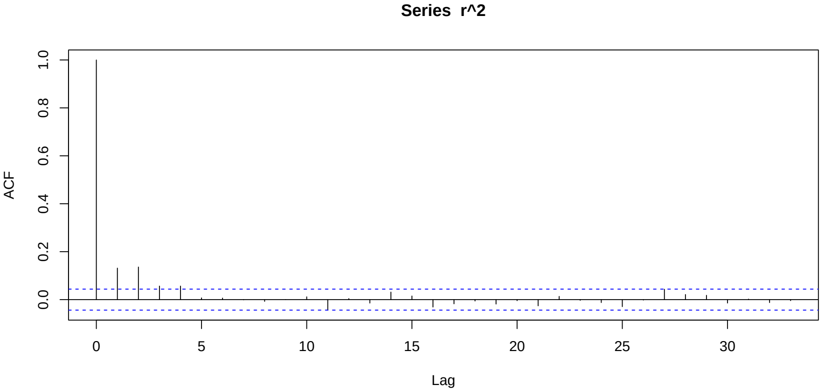
La fonction d'autocorrélation de r_t artificiel

```
acf(r)
```



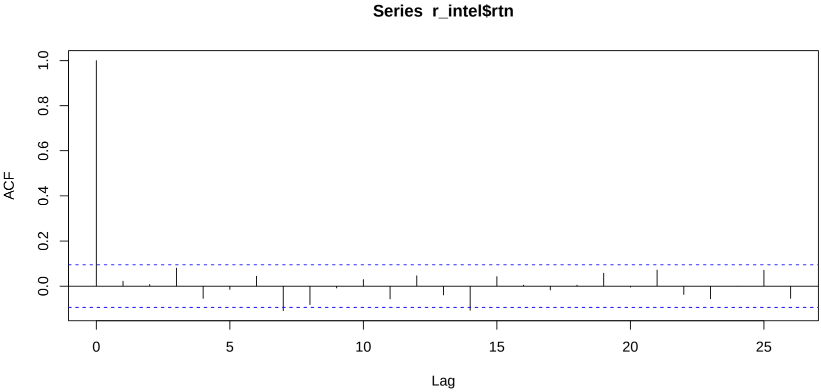
La fonction d'autocorrélation de r_t^2 artificiel

```
acf(r^2)
```



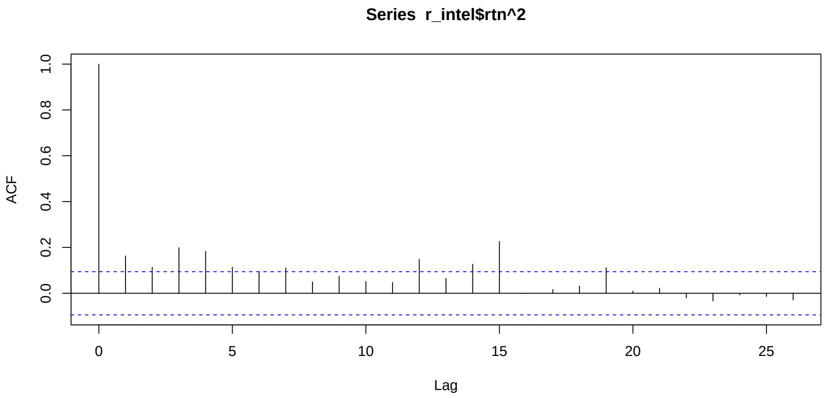
La fonction d'autocorrélation de r_t, Intel 1973-2008

```r
r_intel = read.table('m-intc7308.txt', header=TRUE)
acf(r_intel$rtn)
```



La fonction d'autocorrélation de r_t^2, Intel 1973-2008

```
acf(r_intel$rtn^2)
```



Test des effets ARCH, Intel

```
Box.test(r_intel$rtn, lag=12)
```

```
##
##  Box-Pierce test
##
## data:  r_intel$rtn
## X-squared = 15.987, df = 12, p-value = 0.1918
```

```
Box.test(r_intel$rtn^2, lag=12)
```

```
##
##  Box-Pierce test
##
## data:  r_intel$rtn^2
## X-squared = 78.075, df = 12, p-value = 9.599e-12
```

La log-vraisemblance pour un ARCH(1) gaussien, $\mu_t = \mu$

- En générale, pour une séries r_t, la log-vraisemblance est

$$L(\theta; r_1, \ldots, r_T) = \sum_{t=1}^{T} \log[f(r_t|\theta, r_1, \ldots, r_{t-1})]$$

- La densité d'une aléa $N(\mu, \sigma^2)$:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}.$$

- La log-densité :

$$\log f(x; \mu, \sigma) = -\frac{1}{2}[\log \sigma^2 + \log 2\pi + (x-\mu)^2/\sigma^2]$$

La log-vraisemblance pour un ARCH(1) gaussien, $\mu_t = \mu$

- Terme $t = 1$:

$$\log f(r_1) = -\frac{1}{2}\left[\log\frac{\alpha_0}{1-\alpha_1} + \log 2\pi + \frac{(r_t-\mu)^2}{\alpha_0/(1-\alpha_1)}\right].$$

- Termes $t > 1$:

$$\log f(r_t|r_{t-1}) = -\frac{1}{2}\left[\log[\alpha_0 + \alpha_1 a_{t-1}^2] + \log 2\pi + \frac{(r_t-\mu)^2}{\alpha_0 + \alpha_1 a_{t-1}^2}\right].$$

- La somme de tous les termes donne $L(\mu, \alpha_0, \alpha_1; r_1, \ldots, r_T)$
- Les “constantes” de normalisation importent. Même si elles ne dépendent pas des r_t actuelles, elles dépendent des paramètres et des rendements retardés.
- Pourquoi en logs?

Évaluation de la vraisemblance

```r
vraisemblance <- function(mu, al0, al1, r) {
  T = length(r)
  a = rep(0, T); # Mémoire réservé
  mu_t = mu
  sigma2_t = al0/(1-al1)
  a[1] = r[1]-mu_t
  L = -0.5*(sigma2_t + log(2*pi) + a[1]^2/sigma2_t)
  for (t in 2:T) {
    mu_t = mu
    sigma2_t = al0 + al1 * a[t-1]^2
    a[t] = r[t] - mu_t
    L = L - 0.5*(sigma2_t + log(2*pi) + a[t]^2/sigma2_t)
  }
}
```

Prévision avec un ARCH(m)

- La meilleure prévision ponctuelle, en termes de perte quadratique, est la moyenne.

- Prévision de σ^2_{t+1} à t :

$$\sigma^2_t(1) \equiv E[\sigma^2_{t+1}|F_t] = \alpha_0 + \alpha_1 a^2_t + \ldots + \alpha_m a^2_{t+1-m}.$$

- Prévision de σ^2_{t+2} à t :

$$\sigma^2_t(2) \equiv E[\sigma^2_{t+2}|F_t] = E[E[\sigma^2_{t+2}|F_{t+1}]|F_t]$$

$$\sigma^2_t(2) = E[\alpha_0 + \alpha_1 a^2_{t+1} + \ldots + \alpha_m a^2_{t+2-m}|F_t]$$

$$\sigma^2_t(2) = \alpha_0 + \alpha_1 \sigma^2_t(1) + \alpha_2 a^2_t + \ldots + \alpha_m a^2_{t+2-m}.$$

- Prévision de σ^2_{t+h} à t :

$$\sigma^2_t(h) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma^2_t(h-i),$$

où $\sigma^2_t(h-i) = a^2_{t+h-i}$ si $h - i \leq 0$ $(t + h - i \leq t)$

Notes sur la prévision ponctuelle

- Il faut avoir plus de structure pour évaluer l'incertitude associée à σ_{t+h}. (Par exemple, on peut spécifier une loi pour ϵ_t)
- L'incertitude concernant les paramètres et le modèle n'est pas pris en compte.

Cours 5, la semaine prochaine

Plan préliminaire

1. La théorie des estimateurs maximum de vraisemblance
2. Évaluation de la vraisemblance des modèles GARCH
3. Modèle EGARCH et l'effet de levier