ECN 6578A, Économétrie des marchés financiers, Hiver 2020

Cours 8 et 9

William McCausland

2020-03-28

Le facteur d'actualisation dans un monde sans risque

- Supposons un monde sans risque, avec taux d'intérêt R_t.
- La valeur à t d'un dollar un an plus tard est de

$$M_{t+1} = \frac{1}{1 + R_{t+1}}.$$

- On appelle M_{t+1} le facteur d'actualisation.
- Il actualise (donne une valeur à t à) un paiement à $t + 1$.
- L'absence d'arbitrage entraîne, pour chaque actif i,

$$P_{it} = P_{i,t+1} M_{t+1}.$$

- En termes équivalents,

$$\frac{P_{i,t+1}}{P_{it}} M_{t+1} = (1 + R_{i,t+1}) M_{t+1} = 1.$$

Le facteur d'actualisation stochastique (FAS)

- Le FAS M_t vérifie pour tout actif (ou portefeuille) i :

$$P_{it} = E_t[P_{i,t+1} M_{t+1}], \quad E_t[(1 + R_{i,t+1}) M_{t+1}] = 1.$$

- Notez bien que M_t ne dépend pas de i.
- Version inconditionnelle (prendre l'espérance des deux côtés) :

$$E[(1 + R_{i,t+1}) M_{t+1}] = 1.$$

- Avancée :

$$E[(1 + R_{it}) M_t] = 1.$$

- Avec $\mathrm{cov}[X, Y] = E[XY] - \mu_X \mu_Y$, on obtient

$$1 = \mathrm{cov}[R_{it}, M_t] + E[1 + R_{it}] E[M_t]$$

$$E[1 + R_{it}] = \frac{1}{E[M_t]} (1 - \mathrm{cov}[R_{it}, M_t]). \tag{1}$$

Le rendement zéro-bêta

- Un actif à rendement R_{ot} est un actif *zéro-bêta inconditionnel* si $\mathrm{cov}[R_{ot}, M_t] = 0$.
- Pour un tel actif,

$$E[1 + R_{ot}] = \frac{1}{E[M_t]},$$

et on obtient (soustraire cette équation de (1))

$$E[R_{it} - R_{ot}] = -E[1 + R_{ot}]\mathrm{cov}[R_{it}, M_t].$$

- Remarquez qu'un actif sans risque est toujours un actif zéro-bêta inconditionnel. (Une constante est non-corrélée avec n'importe quelle v.a.)

Deux approches à la dérivation du FAS

- maximisation de l'utilité (hypothèse plus forte), et
- absence de l'arbitrage (moins forte).

Maximisation d'utilité espérée intertemporelle

- Voici une fonction d'utilité sur les chemins aléatoires $\{C_\tau\}_{\tau=t}^{\infty}$ de la consommation :

$$V = E_t\left[\sum_{j=0}^{\infty} \delta^j U(C_{t+j})\right].$$

- V est additivement séparable en temps,
- V est additivement séparable en état (utilité espérée)
- Le taux de préférence de temps (δ) est constant.
- La maximisation de V sous des contraints (richesse, actifs disponibles) donne plusieurs conditions de première ordre, dont

$$U'(C_t) = \delta E_t[(1 + R_{i,t+1})U'(C_{t+1})],$$

  où $R_{i,t+1}$ est le rendement de l'actif i à $t+1$.
- C'est une équation dite d'Euler.

Maximisation d'utilité et le FAS

- Le FAS est le taux marginal inter-temporel de substitution:

$$M_{t+1} = \delta U'(C_{t+1})/U'(C_t),$$

  qui est toujours positif.
- Intuition pour

$$E[R_{it} - R_{ot}] = -E[1 + R_{ot}]\mathrm{cov}[R_{it}, M_t] :$$

  - Si $\mathrm{cov}[R_{it}, M_t] > 0$, i a une valeur relativement élevée quand la consommation future est plus valorisée (i.e. quand $U'(C_{t+1})$ est élevé).
  - En équilibre, son prix est plus élevé (par rapport a un actif j où $\mathrm{cov}[R_{jt} M_t] < 0$) et son rendement moyen est moins élevé.
  - L'investisseur supporte ce rendement moyen moins élevé car l'actif paie quand la consommation est plus valorisée.

Absense d'arbitrage et le FAS

- Voici un milieu très simple :
  - Il y a deux périodes.
  - L'état du monde est aléatoire dans la deuxième période.
  - Il y a S états du monde possible : $1, \ldots, S$.
  - Chaque état s a une probabilités π_s d'être réalisé.
  - Il y a N actifs, chacun avec un paiement dans le deuxième période qui dépend de s.
  - Actif i a un prix q_i en période 1.
  - Actif i paie X_{si} si l'état s se produit.
  - π, X et q sont primitifs.
- Un arbitrage est un portefeuille ω tel que
  - $\omega^\top q \leq 0$ (on ne paie rien dans la première période)
  - $X\omega \geq 0$ (on ne peut pas perdre dans la deuxième), et
  - $X\omega \neq 0$ (on gagne dans au moins un état du monde).

Prix et rendements

- Vecteur q donne les prix à période 1 :

$$\underset{N\times 1}{q} = \begin{bmatrix} q_1 \\ \vdots \\ q_N \end{bmatrix}.$$

- Matrice X donne les paiements des actif :

$$\underset{S\times N}{X} = \begin{bmatrix} X_{11} & \cdots & X_{1N} \\ \vdots & \ddots & \vdots \\ X_{S1} & \cdots & X_{SN} \end{bmatrix}.$$

- Matrix G donne le rendement brut de l'actif i en état s ($G_{si} = X_{si}/q_i$):

$$\underset{S\times N}{G} = \begin{bmatrix} G_{11} & \cdots & G_{1N} \\ \vdots & \ddots & \vdots \\ G_{S1} & \cdots & G_{SN} \end{bmatrix}.$$

Prix d'états

- L'idée : le prix de l'état s est le prix dans la première période d'un actif qui paie 1 si l'état se produit, 0 autrement.
- Définition : vecteur p $(S \times 1)$ est un vecteur des prix d'états si

$$X'p = q,$$

ou, ce qui est équivalent, pour chaque actif i,

$$q_i = \sum_{s=1}^{S} X_{si} p_s.$$

- Si on divise chaque rangée i de $X'p = q$ par q_i, on obtient

$$G'p = \iota.$$

- Ligne i de cette équation vectorielle, $i = 1, \ldots, N$:

$$1 = \sum_{s=1}^{S} G_{si} p_s = \sum_{s=1}^{S} (1 + R_i) p_s.$$

Implications de l'absence d'arbitrage

- Possible en principle : aucun vecteur p, un p, plusieurs p (une question d'algèbre linéaire)
- Résultat très important : pas d'arbitrage ssi il existe un vecteur p positif des prix d'états.
- Si, en plus, $\mathrm{rang}(X) = S$, le marché est dit complet et le vecteur p est unique.
- On peut définir une variable aléatoire M, qui s'avère être le FAS : dans chaque état s, $M_s = p_s/\pi_s$.
- L'absence d'arbitrage implique qu'il existe un p positif t.q. $G'p = \iota$.
- Donc, $M > 0$ et

$$1 = \sum_{s=1}^{S} p_s(1 + R_{si}) = \sum_{s=1}^{S} \pi_s M_s(1 + R_{si}) = E[(1 + R_i)M].$$

- Le résultat se généralise (plusieurs périodes, temps continu).

Le modèle CCAPM

- L'équation
$$1 = E_t[(1 + R_{i,t+1})M_{t+1}],$$
avec
$$M_{t+1} = \delta \frac{U'(C_{t+1})}{U'(C_t)},$$
où C_t est la consommation agrégée est le modèle CCAPM (C pour consommation).

La fonction d'utilité isoélastique

- La fonction d'utilité isoélastique est

$$U(C_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma}, \quad \gamma > 0.$$

- $\lim_{\gamma \to 1} U(C_t) = \log C_t$
- Dans le contexte où les agents maximisent l'espérance de $U(C_t)$ isoélastique, l'aversion relative pour le risque est constante et égale à γ :

$$-C_t \frac{U''(C_t)}{U'(C_t)} = \gamma$$

- Dans un modèle inter-temporel séparable dans le temps avec l'utilité isoélastique à chaque période, l'élasticité de substitution inter-temporelle est constante, et égale à $\psi = \gamma^{-1}$.

Un problème à deux périodes sans incertitude

- Supposez qu'il y a un rendement R sans risque.

$$\max_{C_t, C_{t+1}} U(C_t) + \delta U(C_{t+1}) \quad \text{tel que} \quad C_t + \frac{1}{1+R} C_{t+1} = m$$

- Conditions nécessaires pour un max:

$$\delta C_{t+1}^{-\gamma} - \frac{1}{1+R} C_t^{-\gamma} = 0 \quad \text{(proportions)}$$

$$C_t + \frac{1}{1+R} C_{t+1} = m \quad \text{(niveaux)}$$

$$\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} = \frac{1}{\delta(1+R)}$$

$$\frac{C_{t+1}}{C_t} = [\delta(1+R)]^{1/\gamma}$$

$$\Delta c_{t+1} = c_{t+1} - c_t \equiv \log \frac{C_{t+1}}{C_t} = \frac{1}{\gamma} \log \delta + \frac{1}{\gamma} \log(1+R)$$

Remarques sur le problème à deux périodes

- $\log \frac{C_{t+1}}{C_t}$ est la log-croissance de la consommation,
- $\frac{1}{\gamma}$ est l'élasticité de substitution inter-temporelle,
- $r = \log(1 + R)$ est le log taux d'intérêt,
- Invariance de l'échelle: C_{t+1}/C_t ne dépend pas de m: $\{C_t\}$ pour un riche est un multiple de celui d'un pauvre.
- Le lien entre le risque et la substitution n'est pas flexible.
- C_t agrégée: si tout le monde a une utilité isoélastique avec le même γ et δ, C agrégée est celle d'un consommateur avec cette utilité ayant la richesse agrégée.
- Une rationalisation du consommateur représentatif.

Tester le CCAPM

- Le CCAPM avec l'utilité isoélastique devient

$$1 = E_t \left[(1 + R_{i,t+1}) \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \tag{2}$$

- Pour tester le CCAPM, on peut estimer γ et tester cette restriction sous une hypothèse supplémentaire : $(1 + R_{it}, C_t)$ est log-normal et homoscédastique.

La loi log-normale

- Définition :

$$X \sim LN(\mu, \sigma^2) \Leftrightarrow \log X \sim N(\mu, \sigma^2)$$

- Moments :

$$\log E[X] = E[\log X] + \frac{1}{2}\mathrm{var}[\log X] = \mu + \sigma^2/2,$$

$$E[X] = e^{\mu+\sigma^2/2}.$$

CCAPM avec log-normalité

- L'équation (2) en logarithmes :

$$0 = \log E_t \left[(1 + R_{i,t+1}) \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right].$$

- Sous l'hypothèse supplémentaire de log-normalité,

$$0 = E_t[r_{i,t+1}] + \log \delta - \gamma E_t[\Delta c_{t+1}] + \frac{1}{2}(\sigma_i^2 + \gamma^2 \sigma_c^2 - 2\gamma \sigma_{ic})$$

- σ_i^2 est la variance de $r_{i,t+1}$,
- σ_c^2 est la variance de Δc_{t+1},
- σ_{ic} est la covariance entre $r_{i,t+1}$ et Δc_{t+1}
- Les variances conditionnelles égalent aux variances inconditionnelles par l'homoscédasticité.

Ajouter un actif sans risque

- S'il existe un actif f sans risque, $\sigma_f^2 = \sigma_{fc} = 0$, et

$$r_{f,t+1} = -\log \delta - \frac{\gamma^2 \sigma_c^2}{2} + \gamma E_t[\Delta c_{t+1}]$$

- Pour un actif arbitraire i,

$$E_t[r_{i,t+1} - r_{f,t+1}] = \gamma \sigma_{ic} - \sigma_i^2/2$$

- Version en rendements simples:

$$\log(E_t[(1 + R_{i,t+1})/(1 + R_{f,t+1})]) = \gamma \sigma_{ic}$$

- La constance de cette prime de risque est une implication malheureuse de l'homoscédasticité, pas cohérente avec les données.

Le casse-tête de la prime des actions (The Equity Premium Puzzle)

- Soit i l'indice S&P500
- Prenez l'effet de commerce (commercial paper) comme proxy pour f.
- Pendant 1889-1994, la moyenne de l'échantillon du R_i est de 6%.
- Celle du rendement log en excès est de 4%.
- C_{t+1}/C_t est très lisse ($\sigma_c = 0.033$),
- Sa covariance avec R_i est très faible ($\sigma_{ic} = 0.0027$).
- Le coefficient de risque nécessaire pour expliquer ces faits est de 19, qui est peu crédible, selon les études micro.

Le mystère du taux sans risque

- Version inconditionnelle de l'expression pour $E_t[r_{f,t+1}]$:

$$E[r_{ft}] = -\log\delta + \gamma g - \frac{\gamma^2\sigma_c^2}{2}$$

  où $g = E[\Delta c_{t+1}]$.
- Les moyennes historiques sont:
  - $E[r_{ft}]$: 1.8%
  - g: 1.8%
  - σ_c^2: 3.3%
- Mais $\gamma = 19$ implique $\delta = 1.12 > 1$.
- Intuition : une grande aversion pour le risque implique un très faible volonté à substituer. Avec $C_{t+1}/C_t > 1$, on a une forte désire à emprunter, qui n'est pas cohérent avec un taux d'intérêt bas et un $\delta < 1$.

Les préférences Epstein-Zin

- Tentative à élucidé le casse-tête de la prime de risque avec des préférences qui brisent le lien entre γ (aversion pour le risque) et ψ (élasticité de substitution inter-temporelle), tout en maintenant l'invariance à l'échelle.
- La définition de l'utilité EZ est récursive:

$$U_t = \left\{(1-\delta)C_t^{(1-\gamma)/\theta} + \delta(E_t[U_{t+1}^{1-\gamma}])^{1/\theta}\right\}^{\theta/(1-\gamma)}$$

où

$$\theta = \frac{1-\gamma}{1-\psi^{-1}}.$$

Utilité isoélastique comme cas spécial

- Pour $\theta = 1$,

$$
\begin{aligned}
U_t^{1-\gamma} &= (1-\delta) C_t^{1-\gamma} + \delta E_t[U_{t+1}^{1-\gamma}] \\
&= (1-\delta) C_t^{1-\gamma} + \delta(1-\delta) E_t C_{t+1}^{1-\gamma} + \delta^2 E_t[U_{t+2}^{1-\gamma}] \\
&= (1-\delta) E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau C_{t+\tau}^{1-\gamma} \right].
\end{aligned}
$$

Équation d'Euler pour les préférences E-Z

- Pour le contraint budgétaire suivant :

$$W_{t+1} = (1 + R_{m,t+1})(W_t - C_t),$$

où W_t est la richesse de l'agent représentatif, y compris le capital humain, et $R_{m,t+1}$ est le rendement du marché, EZ montrent que l'équation d'Euler est

$$1 = E_t \left[\left\{ \delta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi} \right\}^{\theta} \left\{ \frac{1}{(1 + R_{m,t+1})} \right\}^{1-\theta} (1 + R_{i,t+1}) \right].$$

- Notez l'invariance de l'échelle : une équation pour le ratio C_{t+1}/C_t.
- Soit X_{t+1} l'intérieur de l'espérance.

CCAPM avec E-Z et log-normalité

- Avec la log-normalité et l'homoscédasticité, on obtient

$$0 = E_t[\log X_{t+1}] + \frac{1}{2}\mathrm{var}_t[\log X_{t+1}],$$

  où

$$\log X_{t+1} = \theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} - (1-\theta) r_{m,t+1} + r_{i,t+1}.$$

- Avec les cas spéciaux $r_i = r_m$ et $r_i = r_f$ et le cas général $r_i = r_i$, on obtient

$$r_{f,t+1} = -\log \delta + \frac{\theta - 1}{2}\sigma_m^2 - \frac{\theta}{2\psi^2}\sigma_c^2 + \frac{1}{\psi}E_t[\Delta c_{t+1}]$$

$$E_t[r_{i,t+1}] - r_{f,t+1} = \left[\theta \frac{\sigma_{ic}}{\psi} + (1-\theta)\sigma_{im}\right] - \frac{\sigma_i^2}{2}$$

- La prime de risque est une somme pondérée de la covariance de r_i avec Δc_{t+1} et sa covariance avec r_m. Exercice: montrez que $\theta = 1$ donne le CCAPM avec utilité isoélastique et que $\theta = 0$ donne approximativement le CAPM.

Utilité non-séparable

- Une autre tentative à élucidé le casse-tête de la prime de risque implique l'utilité non-séparable, tout en maintenant l'invariance à l'échelle:

$$U_t = \sum_{j=0}^{\infty} \delta^j \frac{(C_{t+j}/X_{t+j})^{1-\gamma} - 1}{1-\gamma},$$

  où

  - Habitude interne: $X_t = C_{t-1}^{\kappa}$ où C_t est la consommation individuelle, ou
  - Habitude externe: $X_t = \bar{C}_{t-1}^{\kappa}$ où C_t est la consommation agrégée.

Habitude interne

- Avec l'habitude interne, la décision C_t a un effet sur X_{t+1}, ce dont le consommateur tient compte :

$$
\begin{aligned}
\frac{\partial U_t}{\partial C_t} &= \frac{\partial}{\partial C_t}\left[\frac{(C_t/C_{t-1}^{\kappa})^{1-\gamma}-1}{1-\gamma}+\delta\frac{(C_{t+1}/C_t^{\kappa})^{1-\gamma}-1}{1-\gamma}\right] \\
&= \left(\frac{C_t}{C_{t-1}^{\kappa}}\right)^{-\gamma}\cdot\frac{1}{C_{t-1}^{\kappa}}+\delta\left(\frac{C_{t+1}}{C_t^{\kappa}}\right)^{-\gamma}\cdot(-\kappa)\cdot C_t^{-\kappa-1}\cdot C_{t+1} \\
&= C_t^{-\gamma}C_{t-1}^{\kappa(\gamma-1)}-\delta\kappa C_t^{\kappa(\gamma-1)}C_{t+1}^{-\gamma}(C_{t+1}/C_t)
\end{aligned}
$$

- Cette utilité marginale dépend de C_{t+1}, qui est aléatoire, et t. L'équation d'Euler est donc

$$
E_t\left[\frac{\partial U_t}{\partial C_t}\right]=\delta E_t\left[(1+R_{t+1})\frac{\partial U_{t+1}}{\partial C_{t+1}}\right].
$$

Habitude externe

- Avec l'habitude externe, la décision C_t d'un individuel n'a aucun effet sur X_t, mais en équilibre, tout le monde prend la même décision :

$$\frac{\partial U_t}{\partial C_t} = \frac{\partial}{\partial C_t}\left[\frac{(C_t/X_t)^{1-\gamma}-1}{1-\gamma}\right]_{X_t=C_{t-1}^{\kappa}} = \left(\frac{C_t}{X_t}\right)^{-\gamma}\cdot\frac{1}{X_t} = C_t^{-\gamma}C_{t-1}^{\kappa(\gamma-1)}$$

$$M_{t+1} = \delta\frac{\partial U_{t+1}/\partial C_{t+1}}{\partial U_t/\partial C_t} = \delta\frac{C_{t+1}^{-\gamma}C_t^{\kappa(\gamma-1)}}{C_t^{-\gamma}C_{t-1}^{\kappa(\gamma-1)}}$$

- Avec log-normalité et homoscédasticité,

$$E_t[r_{i,t+1}-r_{f,t+1}] + \sigma_i^2/2 = \gamma\sigma_{ic}$$

$$r_{f,t+1} = -\log\delta - \gamma^2\sigma_c^2/2 + \gamma E_t[\Delta c_{t+1}] - \kappa(\gamma-1)\Delta C_t$$

- La prime de risque ne change pas.
- Encore seulement une γ très grande peut expliquer les données.
- Mais une telle aversion pour le risque est plus cohérente avec un rendement sans risque plus bas.

Motivation pour la GMM

- GMM, c'est la Méthode de Moments Généralisée
- La log-normalité de $(C_{t+1}, R_{i,t+1})$ est une hypothèse très forte.
- L'approche GMM n'exige, comme modèle, qu'une condition de moment inconditionnel comme $E[(1 + R_{it})M_t] = 1$.
- Rappelons que dans le modèle CCAPM

$$M_{t+1} = \delta \frac{U'(C_{t+1})}{U'(C_t)},$$

et dans le cas spécial d'utilité isoélastique,

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma},$$

donc la *fonction de moment* $(1 + R_{it})M_t$ à l'intérieur de l'espérance est une fonction et des données et des paramètres inconnus.

Moments conditionels et inconditionels

- La version conditionnelle de la condition de moment:

$$E_t[(1 + R_{i,t+1})M_{t+1}] = 1.$$

- La version inconditionelle de la condition de moment:

$$E[(1 + R_{i,t+1})M_{t+1}] = 1.$$

- Problème : la version inconditionnelle est beaucoup moins forte
  - moins d'information pertinente pour estimer les paramètres
  - moins d'implications falsifiable, dans un contexte de test
- Solution : si la version conditionnelle tient et un instrument Z_t est observé à chaque période t,

$$\begin{aligned} E[(1 + R_{i,t+1})M_{t+1}Z_t] &= E[E_t[(1 + R_{i,t+1})M_{t+1}Z_t]] \\ &= E[Z_t E_t[(1 + R_{i,t+1})M_{t+1}]] \\ &= E[Z_t] \end{aligned}$$

- Alors la condition $E[((1 + R_{i,t+1})M_{t+1} - 1)Z_t] = 0$ est une *autre* condition de moment inconditionnelle.

Éléments de la GMM

- Une séries de vecteurs aléatoires : w_t $(J \times 1$ à chaque période t, observé en T périodes)
- Vecteur de paramètres : θ_0 $(P \times 1)$
- Fonction de moment : $g(w_t, \theta_0)$ $(Q \times 1)$
- Condition de moment de la population : $E[g(w_t, \theta_0)] = 0$

Exemple CCAPM 1 : w_t

- Vecteur de variables aléatoires :

$$w_t = (C_t, C_{t+1}, R_{t+1}, Z_t)$$

- C_t est la consommation agrégée (1×1).
- R_{t+1} est un vecteur $N \times 1$ de rendements nets.
- Z_t est un vecteur $K \times 1$ de variables exogènes, ou instruments, connu à t.
- Par exemple,

$$Z_t = (1, C_t/C_{t-1}, C_{t-1}/C_{t-2}, R_t, R_{t-1}).$$

Exemple CCAPM 2 : θ_0

- Vecteur de paramètres :

$$\theta_0 = (\delta_0, \gamma_0)$$

- δ_0 et γ_0 sont des “vraies” valeurs des paramètres d’utilité isoélastique :

$$V = E\left[\sum_{\tau=0}^{\infty} \delta_0^{\tau} \frac{C_\tau^{1-\gamma_0}}{1-\gamma_0}\right]$$

Exemple CCAPM 3 : $g(w_t, \theta_0)$

- Fonction de moment :

$$g(w_t, \theta_0) = [(1 + R_{t+1})\delta_0(C_{t+1}/C_t)^{-\gamma_0} - \iota] \otimes Z_t$$

- ι est un vecteur $N \times 1$ de 1.
- $\otimes$ est l'opération de produit Kronecker.
- $[\ldots]$ est $N \times 1$, Z_t est $K \times 1$ donc le produit Kronecker est $NK \times 1$.
- Un élément de $g(w_t, \theta_0)$ pour chaque combinaison d'actif et d'instrument.

Exemple CCAPM 4 : Condition de moment de la population

- Condition de moment de la population :

$$E[g(w_t, \theta_0)] = 0.$$

- Selon le modèle, pour chaque instrument k et chaque actif i,

$$\begin{aligned}
& E[((1 + R_{i,t+1})\delta_0(C_{t+1}/C_t)^{-\gamma_0} - 1) \cdot Z_{kt}] \\
& \quad = \quad E[E_t[((1 + R_{i,t+1})\delta_0(C_{t+1}/C_t)^{-\gamma_0} - 1) \cdot Z_{kt}]] \\
& \quad = \quad E[E_t[((1 + R_{i,t+1})\delta_0(C_{t+1}/C_t)^{-\gamma_0} - 1)]Z_{kt}] \\
& \quad = \quad E[0 \cdot Z_{kt}] = 0
\end{aligned}$$

- Donc les conditions de moment sont entraînées par la théorie.

Identification et suridentification

- Dans cet exemple, $Q = N \times K$ (dimension de g) et $P = 2$ (dimension de θ_0)
- Si $Q \geq P$ et les instruments sont valides, le système est identifié.
- Si $Q > P$, le système est sur-identifié.

La condition de moment de l'échantillon

- Correspondant à $E[g(w_t, \theta_0)]$, la condition de moment de la population, il y a une condition de moment de l'échantillon :

$$g_T(w, \theta) \equiv \frac{1}{T} \sum_{t=1}^{T} g(w_t, \theta)$$

- $w \equiv (w_1, \ldots, w_T)$.
- Comme $g(w_t, \theta)$, $g_T(w, \theta)$ est $Q \times 1$.
- Intuition pour GMM :
  - Si le modèle est vrai, il devrait y avoir une valeur du paramètre θ pour laquelle $g_T(w, \theta)$ est près de zéro.
  - On estime θ avec la valeur pour laquelle $g_T(w, \theta)$ est le plus près à zéro.
  - $g_T(w, \theta)$ étant vecteur, il faut choisir un distance à zéro.
  - On peut évaluer le modèle par la proximité de $g_T(w, \theta)$ à zéro.

L'estimation GMM

- L'estimation GMM (méthode de moments généralisée) est la valeur $\hat{\theta}_{GMM}$ qui minimise

$$Q_T(\theta) = g_T(w,\theta)'\, W_T\, g_T(w,\theta), \tag{3}$$

  où W_T est une matrice $Q \times Q$ définie positive.
- $W_T = I_Q$ (matrice d'identité) donne la somme des carrées.
- Une condition nécessaire pour une solution $\hat{\theta} = \arg\min_\theta Q_T(\theta)$ est

$$2G_T(\theta)' W_T g_T(w,\theta) = 0, \tag{4}$$

  où

$$G_T(w,\theta) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial g_T(w_t,\theta)}{\partial \theta'}.$$

- On peut minimiser Q_T dans (3) ou résoudre (4).

Le choix de W_T, la matrice de pondération

- La théorie de GMM dit que la W_T optimale est

  $$W_T = S^{-1},$$

  où

  $$S = \mathrm{avar}[T^{1/2} g_T(w, \theta_0)].$$

- Notez que S dépend de θ_0, qui est inconnu.
- Même l'estimation de θ dépend de W_T.
- Cette situation mène à une approche itérative.

Estimation de θ à plusieurs étapes

- Commencez avec $W_T = I_Q$.
- Plus robuste mais un peu plus difficile : commencez avec une matrice diagonale avec les précisions de l'échantillon des éléments de $g_T(w, \theta)$.
- Pour calculer ce dernier, on peut choisir des valeurs raisonnables de θ ($\delta = 0.99$ et $\gamma = 3.0$, par exemple).
- Itérez sur les étapes suivantes :
  - Calculer $\hat{\theta}$ par minimisation ou la solution de (4).
  - Calculer
$$\hat{S}(w, \hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} g(w_t, \hat{\theta}) g(w_t, \hat{\theta})'$$
- Arrêtez quand (par exemple) $\max |\hat{\theta}^K - \hat{\theta}^{K-1}| < \epsilon$.
- Pour $\gamma \approx 3$, $\delta \approx 1$, $\epsilon = 10^{-4}$ est raisonnable.

Propriétés Asymptotiques

- Si $g(w_t, \theta)$ est stationnaire et ergodique et W_T est définie positive,
  - $\hat{\theta} \to \theta_0$ en probabilité et
  - $T^{1/2}(\hat{\theta} - \theta_0)$ converge en loi à $N(0, (G'S^{-1}G)^{-1})$.
- Avec la matrice 2×2 de covariance asymptotique $(G'S^{-1}G)^{-1}$, reportez les écarts types (racines carrées des éléments diagonaux)