Dynamic LQ with Shannon capacity

Christopher A. Sims
Princeton University
sims@princeton.edu

November 2, 2020

Problem statement

We are choosing the joint distributions of the stochastic processes $\{Y_t\}$, $\{X_t\}$, subject to the requirement that X_t be in the information set at t, $\mathcal{I}_t$ and to a cost of mutual information between the two processes. Our choice for the joint distribution is subject to this assumed dynamic relation between X and Y:

$$Y_t = GY_{t-1} + HX_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t \mid \{X_s, Y_s, s < t\} \sim N(0, \Omega)\,. \tag{2}$$

At time t, based on $\mathcal{I}_t$, $Y_t \mid \mathcal{I}_t \sim N(\hat{Y}_t, \Sigma_t)$. Therefore

$$Y_{t+1} \mid \mathcal{I}_t \sim N(G\hat{Y}_t + HX_t, G\Sigma_t G' + \Omega)\,. \tag{3}$$

We collect information at $t+1$ that reduces the variance of our distribution for Y_{t+1}, and we have a no-forgetting constraint:

$$G\Sigma_t G' + \Omega - \Sigma_{t+1} \text{ p.s.d.} \quad . \tag{4}$$

Our objective function is

$$\max E\Big[\sum_{t=0}^{\infty} \beta^t \Big(-Y_t' A Y_t + Y_t' B X_t - X_t' C X_t - \theta\big(\log|G\Sigma_{t-1}G' + \Omega| - \log|\Sigma_t|\big)\Big)\Big] \tag{5}$$

Conditioning on $\mathcal{I}_t$

Let $E[Y_t \mid \mathcal{I}_t] = \hat{Y}_t$. Then, by the law of iterated expectations, we can rewrite the objective function as

$$E\left[\sum_{t=0}^{\infty} \beta^t \Big(\overbrace{-\hat{Y}_t' A \hat{Y}_t - \operatorname{tr}(A\Sigma_t) + \hat{Y}_t B X_t - X_t' C X_t}^{E[-Y_t' A Y_t + Y_t' B X_t - X_t' C X_t | \mathcal{I}_t]} \right.$$
$$\left. - \theta\Big(\log(|G\Sigma_{t-1}G' + \Omega|) - \log|\Sigma_t|\Big)\Big)\right]. \quad (6)$$

The evolution of $\hat{Y}_t$ satisfies

$$\hat{Y}_t = G\hat{Y}_{t-1} + HX_{t-1} + G(Y_{t-1} - \hat{Y}_{t-1}) + \hat{Y}_t - Y_t + \varepsilon_t\,. \quad (7)$$

Certainty equivalent solution for X

If we let the error term in (7) be $\eta_t = G(Y_{t-1} - \hat{Y}_{t-1}) + \hat{Y}_t - Y_t + \varepsilon_t$, then $E[\eta_t \mid \mathcal{I}_{t-1}] = 0$. The optimal solution for the X_t and Σ_t sequences jointly will surely be optimal for X_t *given* the Σ_t sequence, so by certainty equivalence the solution for X_t is the solution to

$$\max_{\hat{Y}} \sum_{t=0}^{\infty} \beta^t (-\hat{Y}_t' A \hat{Y}_t + \hat{Y}_t B X_t - X_t' C X_t) \tag{8}$$

$$\text{subject to} \quad \hat{Y}_t = G\hat{Y}_{t-1} + HX_{t-1}\,. \tag{9}$$

It will make X_t a linear function of $\hat{Y}_t$, $X_t = F\hat{Y}_t$.

Recursive form

To put the problem in recursive form and use the $X_t = F\hat{Y}_t$ result we have already derived, we first re-index the information set. Now $\mathcal{I}_t$ refers to the information set available when Σ_t is chosen, as this is now the only control variable being chosen. Σ_t influences the period objective function at time t, but is chosen before η_{t+1}, and hence $\hat{Y}_t$, is known. So what we are now calling $\mathcal{I}_t$ is what above, in deriving the decision rule for X_t, we called $\mathcal{I}_{t-1}$.

$$\begin{aligned}\hat{\hat{Y}}_t &= E[Y_t \mid \mathcal{I}_t] \\ &= (G+HF)\hat{Y}_{t-1} = (G+HF)\hat{\hat{Y}}_{t-1} + \overbrace{(G+HF)(\hat{Y}_{t-1} - \hat{\hat{Y}}_{t-1})}^{\eta_t}.\end{aligned} \tag{10}$$

$$\begin{aligned}W_t &= \mathrm{Var}(Y_t \mid \mathcal{I}_t) \\ &= (G+HF)(W_{t-1} - \Sigma_{t-1})(G+HF)' + G\Sigma_{t-1}G' + \Omega\,.\end{aligned} \tag{11}$$

Discussion of the dynamic equations

- The control now is Σ_t. It not only enters (11), but also (10), through the fact that $\eta_t = \Lambda_{t-1}\xi_t$, where $\Lambda_{t-1}\Lambda'_{t-1} = W_{t-1} - \Sigma_{t-1}$ and ξ_t is a vector of $N(0, I)$ random variables. η_t is the new information arriving after Σ_t is chosen.

- The three terms on the right of (11) arise from uncertainty about $\hat{Y}_{t-1}$ given information at $t-1$, uncertainty about Y_{t-1} given $\hat{Y}_{t-1}$, and ε_t, respectively.

Bellman equation

The optimal solution for Σ_t as a function of the state, and the value function $V(W_t, \hat{\hat{\hat{Y}}}_t)$ are defined by

$$V(W_t, \hat{\hat{\hat{Y}}}_t) =$$

$$\max_{\Sigma_t} E_t[-\operatorname{tr}(A\Sigma_t) - \hat{Y}_t' A \hat{Y}_t + \hat{Y}_t' BF\hat{Y}_t - \hat{Y}_t' F'CF\hat{Y}_t]$$

$$-\theta(\log|W_t| - \log|\Sigma_t|) + \beta E_t V(W_{t+1}, \hat{\hat{\hat{Y}}}_{t+1}) =$$

$$\max_{\Sigma_t} - \operatorname{tr}(W_t A) - \hat{\hat{Y}}_t'(A - BF + F'CF)\hat{\hat{Y}}_t - \operatorname{tr}\big((-BF + F'CF)(W_t - \Sigma_t)\big)$$

$$-\theta(\log|W_t| - \log|\Sigma_t|) + \beta E_t V(W_{t+1}, \hat{\hat{Y}}_{t+1})\,. \quad (12)$$

subject to (10) and (11) and the no-forgetting constraint.

Comments on Bellman equation

- Σ_t is chosen at t, and thereby determines $\mathrm{Var}(\eta_{t+1})$, but $\hat{Y}_t$ and η_{t+1} are not known when Σ_t is chosen.

- η_{t+1} is the new information that affects outcomes at t and information available at $t+1$.

- The "E_t" notation in the Bellman equation means "expectation conditional on the time-t information set", which includes η_t, but not η_{t+1}.

- This Bellman equation could be iterated to a solution via value function iteration, but in problems with many states this would be slow. It seems likely that it is possible to put more structure on the problem.

Pure tracking

- In the pure tracking case, the terms in $\hat{\hat{Y}}_t$ drop out.
- This happens for example when the period objective is just $E[\|Y_t - X_t\|^2]$ and $H = 0$.
- More generally, the uncertainty, conditional on information at t, about $\hat{\hat{Y}}_t$, complicates evaluation of the $E_t V_{t+1}$ term at the end of (12).