Solving for optimal weight for almost finite dependent estimator

Hiro Kasahara

December 20, 2018

1 Finite dependence

Proposition 1 (Characterization of Finite Dependence)**.** *Suppose that there exists a sequence* $\{d^*_{t+\tau}(z)\}_{\tau=0}^{\rho}$ *such that* $(\tilde{\boldsymbol{f}}_{d^*_t(z),t}(d,z))^\top \prod_{\tau=1}^{\rho} \boldsymbol{F}_{d^*_{t+\tau},t+\tau} = \boldsymbol{0}$ *for all* $(d,z) \in \mathcal{D} \times \mathcal{Z}$*, then the model exhibits the* $(\rho+1)$*-period finite dependence with*

$$\tilde{v}_{d^*_t(z),t}(d,z) = \tilde{u}_{d^*_t(z),t}(d,z) + (\tilde{\boldsymbol{f}}_{d^*_t(z),t}(d,z))^\top \sum_{\tau=1}^{\rho} \beta^\tau \boldsymbol{F}_{d^*_{t+\tau}(z),t+\tau} \left(\nabla \bar{\boldsymbol{e}}^{\boldsymbol{P}_{t+1}} + \boldsymbol{u}_{d^*_{t+\tau}(z)} + \boldsymbol{e}^{\boldsymbol{P}_{t+1}}_{d^*_{t+\tau}(z)} \right).$$

Arcidiacono and Miller (2016)

When the model exhibits finite dependence, given the initial estimate for $\{\boldsymbol{p}_{t+\tau} : \tau = 1, ..., \rho\}$, the conditional choice probabilities can be evaluated without solving the Bellman equation.

[**Add examples for finite dependence here**]

Lemma 1. *The model displays finite dependence if and only if* $\left(\tilde{\boldsymbol{F}}^T \otimes \tilde{\boldsymbol{F}}\right) vec(\tilde{\boldsymbol{P}}) = vec(-\tilde{\boldsymbol{F}}\boldsymbol{F_0})$ *has a solution* $\tilde{\boldsymbol{P}}$*.*

Proof. Suppose the model displays finite dependence property, for each $x^{(i)}$, we have

$$\left(\boldsymbol{P_t(x^{(i)})} - \boldsymbol{P'_t(x^{(i)})}\right) \boldsymbol{F(x^{(i)})}\,(\boldsymbol{P_{t+1}F}) = \boldsymbol{0}.$$

Define $\mathcal{K} = \{\boldsymbol{D} \in \mathbb{R}^{D+1}, \sum_i \boldsymbol{D}_i = 0\}$ as the linear space of vectors that sum up to 0. As $\boldsymbol{P_t}$ and $\boldsymbol{P'_t}$ are arbitrary, we have that

$$\boldsymbol{DF(x^{(i)})}\,(\boldsymbol{P_{t+1}F}) = \boldsymbol{0}.\ \forall\ \boldsymbol{P} \in \mathcal{S}$$

With $\boldsymbol{PF} \in \mathcal{K}$ as well, we have that $rank(\boldsymbol{P}_{t+1}\boldsymbol{F}) = 1$, and $dim(null(\boldsymbol{P}_{t+1}\boldsymbol{F})) = |\mathcal{X}| - 1$. From the above results, we have that

$$\tilde{\boldsymbol{F}}\left(\boldsymbol{P_{t+1}}\boldsymbol{F}\right) = \boldsymbol{0}$$
$$\tilde{\boldsymbol{F}}\left(\tilde{\boldsymbol{P}}_{\boldsymbol{t+1}}\tilde{\boldsymbol{F}} + \boldsymbol{F_0}\right) = \boldsymbol{0}.$$

Let $A \otimes B$ be the Kronecker product of two matrix A and B and $\text{vec}(A)$ be the vectorization of matrix A. Then by the property of vectorization

$$\left(\tilde{\boldsymbol{F}}^T \otimes \tilde{\boldsymbol{F}}\right)\text{vec}(\tilde{\boldsymbol{P}}_{\boldsymbol{t+1}}) = \text{vec}(-\tilde{\boldsymbol{F}}\boldsymbol{F_0}).$$

Thus the model displays finite dependence property if and only if above equation has a suitable solution for $\boldsymbol{P_{t+1}}$. □

To find the optimal weight that minimizes $\left(\tilde{\boldsymbol{F}}^T \otimes \tilde{\boldsymbol{F}}\right)(\tilde{\boldsymbol{P}}_{\boldsymbol{t+1}}) + \text{vec}(\tilde{\boldsymbol{F}}\boldsymbol{F_0})$, we solve the linear equation of $\left(\tilde{\boldsymbol{F}}^T \otimes \tilde{\boldsymbol{F}}\right)\text{vec}(\tilde{\boldsymbol{P}}_{\boldsymbol{t+1}}) = \text{vec}(-\tilde{\boldsymbol{F}}\boldsymbol{F_0})$.

Note that $vec(\tilde{\boldsymbol{P}})$ has many zeros, then we need to select the non-zero rows that corresponds to the conditional choice probabilities.

A conditional choice probabilities $\tilde{\boldsymbol{P}} \in \mathcal{P}$ is

$$\tilde{\boldsymbol{P}} = \begin{bmatrix} p(1,x^{(1)}) & \dots & p(D,x^{(1)}) & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & p(1,x^{(|\mathcal{X}|)}) & \dots & p(D,x^{(|\mathcal{X}|)}) \end{bmatrix}_{|\mathcal{X}|\times|\mathcal{X}|D}$$

(According to wikipedia), if $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, then $vec(A) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$. The rows that are non-zero in $\tilde{\boldsymbol{P}}$ are $\{k * |D\mathcal{X}| + m|\mathcal{X}| + k + 1\}$ where $k = 0, \dots, D$,$m = 0, \dots, D$.

Then we can compute the conditional choice probability with OLS using the selected rows. However, this may suffer from the fact that, the parameters are not bounded between 0 and 1.

References

Arcidiacono, P. and Miller, R. A. (2016). Nonstationary Dynamic Models with Finite Dependence *.