Modelos Dinâmicos I: "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher" por John Rust

June 10, 2020

Outline

Introduction

- Estimação de modelos de otimização dinâmica.
- Modelo regenerativo de decisão ótima de parada para descrever Harold Zurcher.
- Trade off entre minimizar custo de manutenção dos ônibus da frota e o custo de falhas em serviço.

Introduction

- Zurcher: superintendente de manutenção da frota de ônibus de Madisin, Wisconsin, EUA.
- Abordagem botton-up para modelar investimento.
- Método: nested fixed point algorithm para estimação.

Base de Dados

- Frota de ônibus de Madison, Wisconsin, EUA.
- Dados de 162 ônibus, dezembro/1975 à maio/1985.
- Para cada ônibus observa todo mês:
  1. milhagem
  2. diário de manutenção. considera retifica do motor somente.

Base de Dados

- Tabela 1: frota de ônibus. Usa grupos 1-4 na estimação.
- Tabela 2a: sub-amostra com pelo menos uma retífica (viés de seleção).
- Tabela 2b: sub-amostra com nenhuma retífica (censura).
- Figura 1: chave do problema.
- Tabela 3: custos de retífica.

Pausa para reflexão

- O que o artigo quer? Descreva o problema a ser resolvido em uma frase.
- O que os dados dizem sobre isso? (ver fig. 1)
- Como você estimaria isso da maneira mais simples possível?
- Essa estimação estaria contemplando todos os aspectos do problema? Se não, qual seria o viés?

Modelo - definições

- variável de controle

$$i_t = \begin{cases} 1, & \text{se houve retífica} \\ 0, & \text{caso contrário} \end{cases}$$

- $x_t \rightarrow$ milhagem em t
- $\{i_t, x_t\} \rightarrow$ processo estocástico controlado
- Odômetro zera quando motor é retificado.

Modelo com solução com forma funcional fechada

$c(x_t, \theta_1) \rightarrow$ custo operacional

Todo mês Zurcher decide entre:

1. manutenção regular com custo $c(x_t, \theta_1)$
2. substitui o motor velho por um novo. Vende o velho por $\underline{P}$, compra um novo por $\overline{P}$, e tem custo operacional $c(0, \theta_1)$.

Modelo com solução com forma funcional fechada

- Utilidade a cada período:

$$u(x_t, i_t) = \begin{cases} -c(x_t, \theta_1), & i_t = 0 \\ -[\overline{P} - \underline{P} + c(0, \theta_1)], & i_t = 1 \end{cases}$$

- Milhagem acumulada a cada mês por um ônibus é distribuido exponencialmente com parâmetro θ_2:

$$p(x_{t+1}|x_t, i_t, \theta_2) = \begin{cases} \theta_2 exp\{\theta_2(x_{t+1} - x_t), & \text{se } i_t = 0 \text{ e } x_{t+1} \geq x_t \\ \theta_2 exp\{\theta_2(x_{t+1}), & \text{se } i_t = 1 \text{ e } x_{t+1} \geq 0 \\ 0, & \text{nos outros casos} \end{cases}$$

Optimal Stopping Problem

- A regra ótima de quando parar é a solução do seguinte problema:

$$V_\theta(x) = \sup_{\Pi} E\{\sum_{j=t}^{\inf} \beta^{j-t} u(x_j, f_j, \theta_1)|x_t\}$$

Onde Π é uma sequência infinita de regras de decisão $\Pi = \{f_t, f_t + 1, ...\}$.

- $V_\theta(x)$ é a função valor e é a única solução da equação de Bellman:

$$V_\theta(x_t) = \max_{i_t} u(x_t, i_t, \theta_1) + \beta EV_\theta(x_t, i_t)\}$$

onde

$$EV_\theta(x_t, i_t) = \int_0^\infty V_\theta(y) p(dy|x_t, i_t, \theta_2)$$

Optimal Stopping Problem

- Rust prova que a solução ótima é uma regra markoviana estacionária do tipo:

$$i_t = f(x_t, \theta) = \begin{cases} 1, & \text{se } x_t > \gamma(\theta) \\ 0, & \text{se } x_t \leq \gamma(\theta) \end{cases}$$

- $\gamma(\theta)$: valor crítico da milhagem. (ver figura 1. qual o problema?)

Modelo estrutural sem forma funcional fechada

- Utilidade: $u(x_t, i_t, \theta_1) + \epsilon_t(i)$
- probabilidad de transição entre estados: $p(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t, i_t, \theta_2, \theta_3)$
- Equação de Bellman:

$$V_\theta(x_t) = \max_{i_t}\{u(x_t, i_t, \theta_1) + \epsilon_t(i) + \beta EV_\theta(x_t, \epsilon_t, i_t)\}$$

onde

$$EV_\theta(x_t, i_t) = \int_0^\infty V_\theta(y, \eta) p(dy, d\eta | x_t, \epsilon_t, i_t, \theta_2, \theta_3)$$

A solução do modelo e a integral da esperança

- CPO de eq. de Bellman fornece a solução ótima.
- Regra de decisão estacionária: $i_t = f(x_t, \epsilon_t, \theta)$
- Problema: ϵ_t aparece em EV_θ. Logo, é necessário calcular a esperança de integrando ϵ. Temos então que integrar V_θ usando $p(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t, i_t, \theta_2, \theta_3)$. Feito isto, temos que integrar a equação de Bellman para achar a política ótima $i_t = f(x_t, \epsilon_t, \theta)$.
- **Note o seguinte: $i(x, \epsilon, \theta)$. Logo, existem probabilidades associadas a cada possível decisão. Em econometria não chamamos de política ótima, mas de conditional choice probabilities.**

Hipótese 1: independência condicional

- $p(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t, i_t, \theta_2, \theta_3) = q(\epsilon_{t+1} | x_{t+1}, \theta_2) p(x_{t+1} | x_t, i_t, \theta_3)$

- x_{t+1} is a sufficient statistic for ϵ_{t+1}
- x_{t+1} depende somente de x_t, e não de ϵ_t

Hipótese 2: extreme value distribution

- $q(\epsilon|x,\theta)$ segue uma distribuição extreme value type I.
- E voltamos pro logit! A probabilidade do agente escolher $i = 1$, $P(i = 1|x,\theta)$ é:

$$P(u(x_t,1,\theta_1)+\epsilon_t(1)+\beta EV_\theta(x_t,\epsilon_t,1) \geq u(x_t,0,\theta_1)+\epsilon_t(0)+\beta EV_\theta(x_t,\epsilon_t,0)) =$$

$$P(\epsilon_t(1)-\epsilon_t(0)) \geq u(x_t,0,\theta_1)-u(x_t,1,\theta_1)+\beta EV_\theta(x_t,\epsilon_t,0)-\beta EV_\theta(x_t,\epsilon_t,1))$$

- Como ϵ é distribuído como valor extremo, a solução da integral para calcular essa probabilidade, tem a forma funcional do logit.

Conditional choice probability

- Conditional choice probability:

$$P(i = 1|x, \theta) = \frac{exp\{u(x_t, 1, \theta_1) + \beta EV_\theta(x_t, \epsilon_t, 1)\}}{\sum_j exp\{u(x_t, j, \theta_1) + \beta EV_\theta(x_t, \epsilon_t, j)\}}$$

onde

$$EV_\theta(x, i) = \int_y log \sum_j exp[u(y, j, \theta_1) + \beta EV_\theta(y, j)]p(dy|x, i, \theta_3)$$

- Esse termo é chamado de excedente social, e também é conhecido como log-sum.
- Paralelo com o nested logit. $EV_\theta(x_t, \epsilon_t, i)$ é o inclusive value do nested logit.

Função verossimilhança

- A funcão verossimilhança para um ônibus é dada por

$$l(x_1, ..., x_T, i_1, ..., i_T | x_0, i_0, \theta) = \prod_{t=1}^{T} P(i_t | x_t, \theta) p(x_t | x_{t-1}, i_{t-1}, \theta_3)$$

Algorítmo: nested fixed point

1. Dado um θ, itera a equação do excedente social para achar o ponto fixo EV_θ (algoritmo interno)
2. Dado EV_θ, algoritmo de busca acha o θ que maximiza a função verossimilhança (algoritmo externo)

Estimando o modelo - parametrizando

- utilidade:

$$u(x_t, i_t, \theta_1) + \epsilon_t(i) = \begin{cases} -c(x_t, \theta_1) + \epsilon_t(0), & \text{se } i_t = 0 \\ -RC - c(0, \theta_1) + \epsilon_t(1), & \text{se } i_t = 1 \end{cases}$$

- probabilidade de transição

$$p(x_{t+1} | x_t, i_t, \theta_3) = \begin{cases} g(x_{t+1} - x_t, \theta_3), & \text{se } i_t = 0 \\ g(x_{t+1} - 0, \theta_3), & \text{se } i_t = 1 \end{cases}$$

Estimando o modelo - parametrizando

- Objetivo é estimar $\beta, \theta_1, \theta_3, RC$

1. discretiza o espaço de estado x_t (milhagem): 90 intervalos de 5000 milhas.
2. $g(.)$ é uma distribuição multinomial sobre $1, 2, 3$, que correspondem aos aumentos de milhas $[0, 5000]$, $[5000, 10000)$ e $[10000, \infty)$. $f(x_1, ..., x_n) = p_1^{x_1} ... p_n^{x_n}$
3. Custo operacional

$$c(x, \theta_1) = \begin{cases} \theta_{11}x + \theta_{12}x^2 + \theta_{13}x^3 \\ \theta_{11}exp(\theta_{12}x) \\ etc \end{cases}$$

4. $\epsilon_t(0), \epsilon_t(1)$ é iid valor extremo

Estimação

- 1o. estágio: estimar a probabilidade de transição entre os estados, $p(x_{t+1}|x_t, i_t, \theta_3)$
- 2o. estágio: usa $\hat{\theta}_3$ e estima $P(i_t|x_t, \theta)$.
- Estima a full likelihood

Resultados

- Tabelas V e VI: resultados do 1o. estágio.
- Estima grupos 1, 2 e 3 juntos.
- Tabela VII: milhagem segue um processo de markov?
- Tabela VIII: procurando a melhor especificação. Escolhe grupos 1, 2 e 3, e custo linear e raiz quadrada.

Resultados - otimizador versus míope

- Tabela IX: parâmetros estruturais. Usa $\beta = 0.99$ e $\beta = 0$ (modelo míope).
- Note a diferença no θ_{11}.
- Modelo míope troca o motor somente quando $c(x, \theta_{11}) > RC + c(0, \theta_{11})$. Entende que o custo cresce mais rápido com x_t.

Resultados - custos

- Custo de retífica para os grupos 1, 2 e 3 é \$9499 (tabela 3). Custo estimado é $RC = 11.72$.
- Relação entre real e estimado é $\sigma = \frac{9499}{11.72}$
- $c(x) = 0.001\theta_{11}x$
- custo estimado é $\sigma * 0.001 * 4.826 = 3.75$ dolares em manutenção por cada 5000 milhas para os ônibus dos grupos 1, 2 e 3.
- um ônibus com 300,000 milhas tem um custo mensal de 225 dólares maior que um 0 milhas.
- esse custo é 1.70 dolares para ônibus do grupo 4. Isso explica porque os ônibus mais novos dos grupos 1, 2 e 3 retificam o motor mais cedo.