Integrated learning and decision making

Stella Kapodistria

joint work with Collin Drent, Paulo Serra

Department of Mathematics, TU/e
My goal: Given a trove of data, be able to choose when to perform action $A$ vs action $B$ (at costs $c_A$ vs $c_B$)
Data \rightarrow \text{Estimation} \rightarrow \text{Statistical model} \rightarrow \text{Optimization} \rightarrow \text{Cost optimal decision}

\{X(t), t \geq 0\}

If $X(t) \leq \tau_t$, then do PM.
If $X(t) \leq 0$, then do CM.
Upon failure, it can be replaced with an as-good-as new iid component at cost $c_c$. Before failure, the component can be preventively replaced at cost $c_p < c_c$.

**Data = Lifetimes**

**Age replacement policy**

RUL of component

- $\tau^*$
- $L$
- time

**Data = Condition**

**Condition replacement policy**

condition of asset $X(\tau) \approx \sigma BM + CPP^\varepsilon$

- $\xi$
- $\tau^*$
- $L_{\xi}$
- time
Condition prescriptive maintenance policy for a heterogeneous population

Traditional condition replacement policy

condition ($\theta$)

$\theta$

Known from historical data

Integrated condition replacement policy

condition ($\theta$)

Updated parameter prior

$p(\theta|x) \propto L(x|\theta)p(\theta)$

Observed data added

Surrogate model
Condition maintenance policy: continuous time

\{X(t) = \alpha + \beta t + \sigma \varepsilon(t), \ t \geq 0\} denotes the condition, with \(\varepsilon(t)\) a standard Brownian motion

Failure is defined as the first passage time of \(X(t)\) to \(\xi\)

Preventive replacement cost = \(c_p = c < c_c = c + z\) = Corrective replacement cost

Question:

What is the optimal preventive maintenance (replacement) policy?

Optimal policy (no information): \(\tau^* = \arg \min_{\tau} \frac{\mathbb{E}[\text{cost in } L_\xi \land \tau]}{\mathbb{E}[L_\xi \land \tau]} = \infty\)

Optimal policy (full information \(\@ \ dt \to 0^+)\): 0^+

Optimal policy (full information \(\@ \ dt > 0\))?
Condition prescriptive policy: discrete time

\( \{X(t_n) = \alpha + \beta t_n + \sigma \varepsilon(t_n), n \geq 0\} \) denotes the condition, with \( t_n - t_{n-1} = \Delta \)

Failure is defined as the first passage time of \( X(t_n) \) to \( \xi \)

Preventive replacement cost = \( c_p = c < c_c = c + z = \) Corrective replacement cost

**Question:**

What is the optimal preventive maintenance (replacement) policy?

Optimal Policy Theorem [K, Serra 2018]: if \( c < (c + z)e^{-r \Delta} \), then there exists an optimal stationary policy \( \tau \).

\[ V_\tau = \mathbb{E}[e^{-rL_\tau}](c + V_\tau) + \mathbb{E}[e^{-rL_\tau}I\{\xi - X(L_\tau) \leq 0\}]z \]

\[ \Rightarrow V_\tau = \frac{\mathbb{E}[e^{-rL_\tau}c + \mathbb{E}[e^{-rL_\tau}I\{\xi - X(L_\tau) \leq 0\}]z]}{1 - \mathbb{E}[e^{-rL_\tau}]} \]
Condition prescriptive policy: discrete time

\{X(t_n) = \alpha + \beta t_n + \sigma \varepsilon(t_n), n \geq 0\} denotes the condition, with \(t_n - t_{n-1} = \Delta\)

Failure is defined as the first passage time of \(X(t_n)\) to \(\xi\)

Preventive replacement cost = \(c_p = c < c_c = c + z\) = Corrective replacement cost

**Question:**

What is the optimal preventive maintenance (replacement) policy?

Optimal Policy Theorem \([K, Serra 2018]\): if \(c < (c + z)e^{-r\Delta}\), then there exists an optimal stationary policy \(\tau\).

\(V_\tau: = \text{total expected discounted cost for the infinite horizon problem}\)

\[
\lim_{r \to 0} r V_\tau = \frac{c + \mathbb{P}[X(L_{\tau}) \geq \xi]z}{\mathbb{E}[L_{\tau}]}
\]

\(\lim_{r \to 0} r V_\tau := \text{long-run rate of cost}\)
Condition prescriptive policy: discrete time

\[{X(t_n)} = \alpha + \beta t_n + \sigma \epsilon(t_n), n \geq 0\] denotes the condition, with \(t_n - t_{n-1} = \Delta\)
Failure is defined as the first passage time of \(X(t_n)\) to \(\xi\)
Preventive replacement cost = \(c_p = c < c_c = c + z\) = Corrective replacement cost

**Question:**

What is the optimal preventive maintenance (replacement) policy?

Optimal Policy Theorem [K, Serra 2018]: if \(c < (c + z)e^{-r\Delta}\), then there exists an optimal stationary policy \(\tau\).

First order approximation: \(\tau \approx \beta \Delta\)

Second order approximation: \(\tau \approx \beta \Delta + \zeta_{c,z} \sigma \sqrt{\Delta}\)
Bayesian condition replacement policy

\( \{X(t_n) = \alpha + \beta t_n + \sigma \varepsilon(t_n), n \geq 0\} \) denotes the condition, with \( t_n - t_{n-1} = \Delta, \alpha, \beta \sim \text{Normal prior} \)

Failure is defined as the first passage time of \( X(t) \) to \( \xi \)

Preventive replacement cost = \( c_p = c < c_c = c + z \) = Corrective replacement cost

**Question:** What is the optimal preventive maintenance (replacement) policy?

**Objective:** determine the (cost) optimal preventive maintenance policy

\[
V(\xi - x, n; \alpha_n, \beta_n) \equiv V(\xi - x, n)
\]

**Approach:** Bellman equations

**Results:** optimal policy = update-dependent control limit policy

\[
V(\xi - x, n) = \begin{cases} 
  c_c + e^{-r} \Delta \mathbb{E}[V(\xi - \alpha_n - \beta_n \delta - \sigma \varepsilon_\delta, n + 1)], & \text{if } \xi - x \leq 0 \\
  \min\{c_p + e^{-r} \Delta \mathbb{E}[V(\xi - \alpha_n - \beta_n \delta - \sigma \varepsilon_\delta, 0), n + 1], e^{-r} \Delta \mathbb{E}[V(\xi - x - \beta_n \delta - \sigma \varepsilon_\delta, n + 1)]\}, & \text{if } \xi - x > 0
\end{cases}
\]

Surrogate model: Bayes condition model
Bayesian condition replacement policy

Design matrix

\[ X_n = T_n \theta + \varepsilon_n, T_n = \begin{bmatrix} 1 & T_1 \\ \vdots & \vdots \\ 1 & T_n \end{bmatrix}, \theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \varepsilon_n \sim N(0, \Sigma_n) \]

\[ T_1 = \Delta, T_{n+1} = (T_n + \delta)I_{\{X_n \leq \tau_n\}} + \delta I_{\{X_n > \tau_n\}} \]

Posterior distribution

\[ \Pi(\alpha, \beta | L_1, ..., L_n) = \text{Normal}\left((I_2 + T'_n \Sigma_n^{-1}T_n)^{-1}T'_n \Sigma_n^{-1}T_n, (I_2 + T'_n \Sigma_n^{-1}T_n)^{-1}\right) \]

Note that

\[ \mathbb{E}\|\hat{\theta}_n - \theta\|^2 \leq \frac{\sigma^2}{n} + \frac{\|\theta\|^2}{n^2} \]

(root-n consistent estimator)
Bayesian condition replacement policy

In the model with known parameters, second order approximation
\[ \tau \approx \beta \Delta + \zeta_{c,z} \sigma \sqrt{\Delta} \]

So we create a regret function that converges to the second order approximation
\[
\mathcal{C}(\tau; x_n, \alpha_n, \beta_n) = c \mathbb{I}_{\{0<\xi-x_{n+1}, \tau<\xi-x_n\}} + (c + z) \mathbb{I}_{\{0<\xi-x_{n+1}, \tau<\xi-x_n\}} \xi + (2c + z) \mathbb{I}_{\{0<\xi-x_{n+1}, \tau<\xi-x_n\}}
\]

We compute the \( \tau \) that minimizes the conditional expectation \( \mathbb{E}[\mathcal{C}(\tau; x_n, \alpha_n, \beta_n)] \). That produces a sequence of \( \tau \)'s, say \{\( \tau_n \)\}.

**Theorem** [K, Serra 2018]:
\[
\tau_n = \beta_n \Delta + \Phi^{-1} \left( \Phi \left( \frac{\xi - \alpha_n - \beta_n \Delta}{\sigma \sqrt{\Delta}} \right) - \frac{c}{c + z} \right) \sigma \sqrt{\Delta}
\]
Policy for detecting failure 20 days prior.
Cost savings in comparison to the oracle

Bayes cost gap < Naïve $\theta$ updating cost gap < Historical $\theta$ cost gap

“Many small jumps”

condition ($\theta$)

“Few large jumps”

condition ($\theta$)
My goal: Given a trove of data, be able to choose when to perform action $A$ vs action $B$ (at costs $c_A$ vs $c_B$)
Integrated learning and decision making

Stella Kapodistria

joint work with Collin Drent, Paulo Serra

Department of Mathematics, TU/e