Model Predictive Controller Tuning by Machine Learning and Ordinal Optimisation

Robert Chin[1,3]

Supervisors: Prof. Chris Manzie[1], Prof. Jonathan E. Rowe[3], Prof. Dragan Nešić[1], Prof. Iman Shames[2]

[1]The University of Melbourne
[2]Australian National University
[3]University of Birmingham

Outline

Introduction

Other Contributions

- Machine Learning Tuning Framework
- Preference Learning
- Active Learning

Ordinal Optimisation

- Lower Bound

Sequential Learning Algorithm for Probabilistically Robust Controller Tuning

- Numerical Example

Further Work

Tuning Model Predictive Controllers

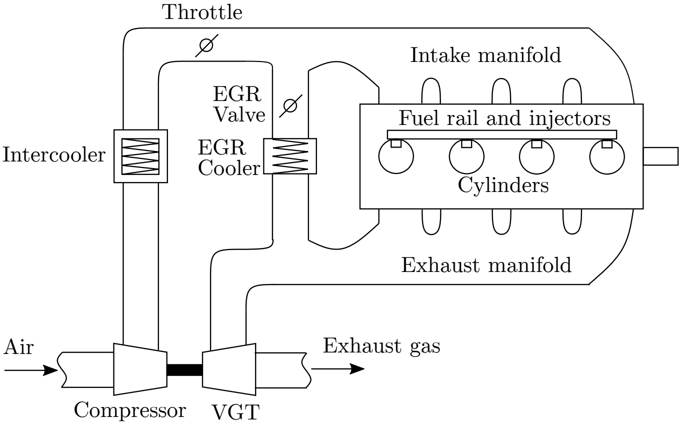
- MPC quadratic cost function:

$$V_k = \sum_{i=0}^{N-1} \left(x_{k|i}^\top Q x_{k|i} + u_{k|i}^\top R u_{k|i} \right) + x_{k|N}^\top P x_{k|N}$$

- The positive definite matrices Q, P, R are tuning variables.
  - Non-trivial relationship with closed-loop trajectory.
- Tuning MPC for performance can be non-intuitive and time-consuming.

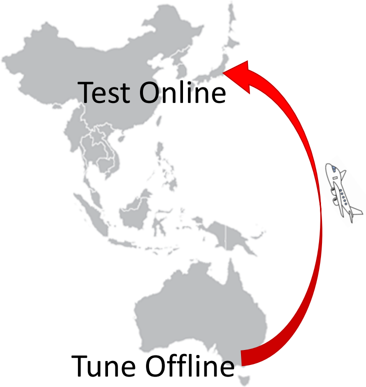
Diesel Air-Path

- Application focus: Automotive diesel engine air-path
- With Toyota Japan



Offline Tuning

- Limited budget for controller testing/tuning on physical plant.
- However, can tune controllers offline in simulation beforehand.



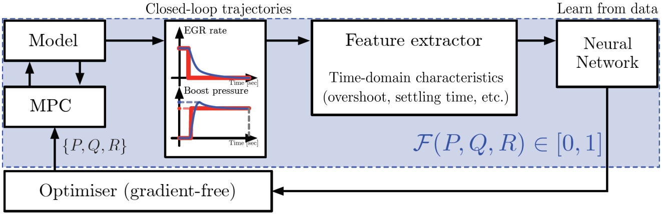
- Algorithms for online MPC tuning successfully demonstrated.[1]
- Want controllers tuned offline to be good initial points

[1] IFAC 2020 (Maass, Manzie, Shames, Chin, Ulapane, Nešić, Nakada)

Outline

Machine Learning Tuning Framework

- Human preferences & trade-offs important in tuning
- Want to replicate preferences in automated offline tuning
- Not easy to write a function describing human preferences
- Proposed framework[2] :

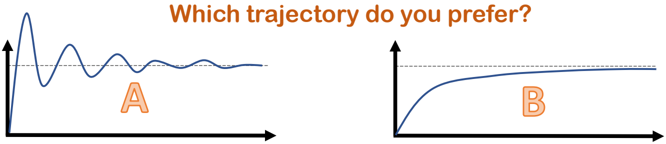


- Learn mapping from time-domain characteristics to scalar performance index.
  - Numeric rating data provided by human experts.

[2] ICARCV 2018 (Ira, Shames, Manzie, Chin, Nešić, Nakada, Sano)

Isotonic Preference Learning from Pairwise Comparisons

- Potential problems with numeric rating labels
  - e.g. scale can 'drift' over time
  - Idea: to solicit pairwise comparisons from experts

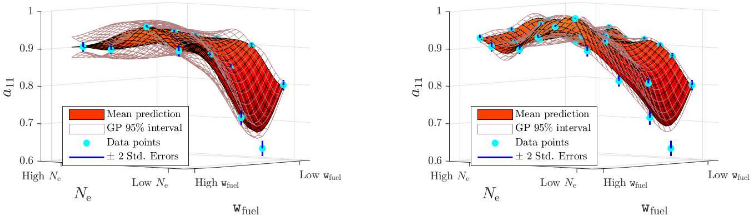


  - Data comprises pairs of features $(\mathrm{x}, \mathrm{x}')$, and binary labels indicating which is preferred.
- Some features may be *desirable*
  - e.g. faster settling time better, all else unchanged.
  - Want monotonicity constraints on particular features (i.e. *isotonic regression*).
- Proposed approach from pairwise comparison data using Gaussian process regression.[3]

[3] CDC 2018 (Chin, Manzie, Ira, Nešić, Shames)

Active Learning for LPV System Identification

- Linear Parameter Varying model suitable for diesel air-path
- Want to identify LPV model by conducting as few experiments as possible
- Active learning approach proposed using Gaussian process regression [4]
  - Conducts next experiment where there is the most uncertainty



- Also used to quantify model uncertainty

[4] IFAC 2020 (Chin, Maass, Ulapane, Manzie, Shames, Nešić, Rowe, Nakada)

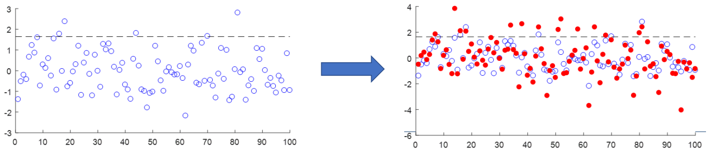
Outline

Ordinal Optimisation and Randomised Algorithms

- Finding approximate solutions to difficult design problems, under uncertainty.
  - OO: from Discrete Event Dynamic Systems literature
  - RA: from Probabilistic Robust Control literature
- Similar philosophy - simulate many random designs and pick the best
- Goal softening to control degree of suboptimality
- Research goal: obtain bound on performance of a controller tuned offline when tested online on a realised plant online.

Ordinal Optimisation Thought Experiment

- 100 job candidates, each with quality drawn i.i.d. $\mathcal{N}(0,1)$
- Interview each candidate, but quality observed with another i.i.d. $\mathcal{N}(0,1)$ perturbation.



- Hire the best 5 observed candidates.
- Of the selected 5, what is the probability that at least one of the hirees is in the top 5% of the $\mathcal{N}(0,1)$ population?

0.1

0.5

0.9

Formalising the Problem

- I.i.d. sample of size n of:

$$\underbrace{Z}_{\text{observation}} = \underbrace{X}_{\substack{\sim\mathcal{N}(0,1) \\ \text{(signal)}}} + \underbrace{Y}_{\substack{\sim\mathcal{N}(0,\xi^2) \\ \text{(noise)}}}$$

- Select the best m observed:

$$\underbrace{Z_{1:n}}_{=X_{\langle 1\rangle}+Y_{\langle 1\rangle}} \leq Z_{2:n} \leq \cdots \leq \underbrace{Z_{m:n}}_{=X_{\langle m\rangle}+Y_{\langle m\rangle}}$$

- Success probability:

$$p_{\text{success}}(n, m, \alpha, \xi) = \Pr\left(\min_{i\in\{1,\ldots,m\}} X_{\langle i\rangle} \leq x^*_\alpha\right)$$

  - where $\alpha > 0$ is the 100α percentile of the distribution of X

Key Observations

$$\underbrace{Z_{1:n}}_{= X_{\langle 1\rangle} + Y_{\langle 1\rangle}} \leq Z_{2:n} \leq \cdots \leq \underbrace{Z_{m:n}}_{= X_{\langle m\rangle} + Y_{\langle m\rangle}}$$

$$p_{\text{success}}(n, m, \alpha, \xi) = \Pr\left(\min_{i \in \{1,\ldots,m\}} X_{\langle i\rangle} \leq x^*_\alpha\right)$$

- What really matters is the joint distribution of (Z, X)
- If each of Z and X are transformed by a strictly increasing function, p_{success} is unchanged
- Can generalise (Z, X) to the class of continuous distributions with a *bivariate Gaussian copula*

Copulas

- Copula: a multivariate distribution with Uniform $(0, 1)$ marginals
- Decouples dependence within a multivariate distribution from its marginals
- Any distribution can be fully represented by its marginals and a copula
  - e.g. to sample from a bivariate distribution, first sample from the copula:
    $$(U_1, U_2)$$
    then generate marginals using inverse probability integral transform:
    $$\left(F_1^{-1}(U_1), F_2^{-1}(U_2)\right)$$
  - Sklar's theorem: for continuous distributions, the copula is unique

- Let
$$\begin{bmatrix} Z \\ X \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$
- Then the bivariate Gaussian copula with correlation ρ is the distribution of:
$$(\Phi(Z), \Phi(X))$$
where Φ is the standard Gaussian CDF

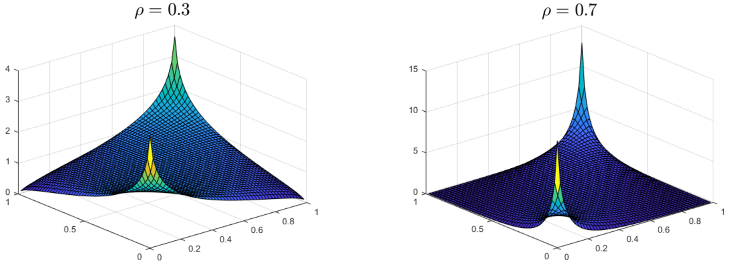


Figure: Bivariate Gaussian copula density

Gaussian Copula OO Success Probability

- Suppose (Z, X) has a continuous distribution with a bivariate Gaussian copula and correlation $\rho > 0$
- Select the best m observed:

$$\underbrace{Z_{1:n}}_{\text{hidden } X_{\langle 1 \rangle}} \leq Z_{2:n} \leq \cdots \leq \underbrace{Z_{m:n}}_{\text{hidden } X_{\langle m \rangle}}$$

- Success probability:

$$p^{\mathcal{N}}_{\text{success}}(n, m, \alpha, \rho) = \Pr\left(\min_{i \in \{1,\ldots,m\}} X_{\langle i \rangle} \leq x^*_\alpha \right)$$

- Equivalence to Gaussian additive noise success probability, with $\rho = \left(1 + \xi^2\right)^{-1/2}$

Properties of Success Probability

1. (Monotonicity) $p^{\mathcal{N}}_{\mathtt{success}}(n, m, \alpha, \rho)$ is non-decreasing in each of n, m, α, ρ
2. (High probability) For any $\delta \in (0, 1]$, can find sufficiently large n so that
$$p^{\mathcal{N}}_{\mathtt{success}}(n, m, \alpha, \rho) \geq 1 - \delta$$
3. (Convergence to 1)
$$\lim_{n \to \infty} p^{\mathcal{N}}_{\mathtt{success}}(n, m, \alpha, \rho) = 1$$

An Analytic Lower Bound

- For any $\omega \in (0, \pi/2)$, let

$$c_1 := \frac{1}{2} - \frac{\omega}{\pi}, \quad c_2 := \frac{\cot \omega}{\pi - 2\omega}$$

$$\mu_n := -\sqrt{\frac{\log(n c_1)}{c_2}}, \quad \sigma_n^2 := \frac{-\log\log 2}{2 c_2 (\log(n c_1) - \log\log 2)}$$

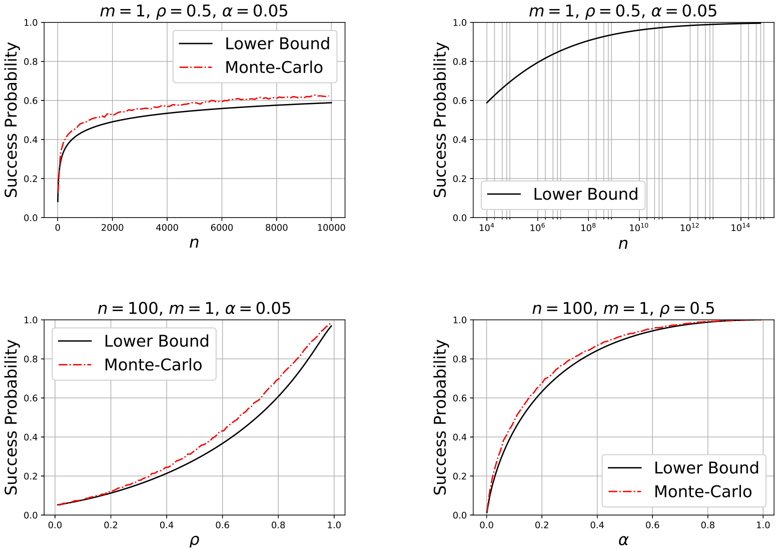
- Then $\exists n^*(\omega) \in \mathbb{N}$ s.t. $\forall n \geq n^*(\omega)$:

$$p_{\text{success}}^{\mathcal{N}}(n, m, \alpha, \rho) \geq \Phi\left(\frac{\Phi^{-1}(\alpha) - \rho\mu_n}{\sqrt{1 - \rho^2 + \rho^2\sigma_n^2}}\right)$$

- Optimised bound w.r.t. ω

$$p_{\text{success}}^{\mathcal{N}}(n, m, \alpha, \rho) \geq \sup_{\omega \in (0, \pi/2): n \geq n^*(\omega)} \Phi\left(\frac{\Phi^{-1}(\alpha) - \rho\mu_n}{\sqrt{1 - \rho^2 + \rho^2\sigma_n^2}}\right)$$

Properties and Lower Bound Illustrated



Inversion of Lower Bound

- The lower bound

$$\underline{p}^{\mathcal{N}}_{\text{success}}(n, m, \alpha, \rho) := \sup_{\omega \in (0, \pi/2): n \geq n^*(\omega)} \Phi\left(\frac{\Phi^{-1}(\alpha) - \rho\mu_n}{\sqrt{1 - \rho^2 + \rho^2\sigma_n^2}}\right)$$

  can be inverted

- For given m, α, ρ and $\delta \in (0, 1]$, find an $\bar{n}$ s.t.

$$p^{\mathcal{N}}_{\text{success}}(n, m, \alpha, \rho) \geq 1 - \delta, \quad \forall n \geq \bar{n}$$

- E.g. with $\delta = 0.01$, $m = 1$, $\alpha = 0.01$, $\rho = 0.9$,

$$\bar{n} = 16744$$

- E.g. with $\delta = 0.01$, $m = 1$, $\alpha = 0.01$, $\rho = 0.01$,

$$\bar{n} \approx 8.144 \times 10^{47007}$$

Non-Gaussian Copula OO Success Probability

- What if (Z, X) is a non-Gaussian copula, but still has 'positive dependence'?
- Denote success probability $p'_{\mathtt{success}}(n, m, \alpha)$
- Properties:
  1. Computed via m-dimensional integral
  2. Non-decreasing in each of n, m, α
  3. General bounds

$$1 - (1-\alpha)^m \leq p'_{\mathtt{success}}(n, m, \alpha) \leq 1 - (1-\alpha)^n$$

  4. In general,

$$\lim_{n \to \infty} p'_{\mathtt{success}}(n, m, \alpha) \neq 1$$

Associated Gaussian Copula

- (Z, X) is an arbitrary copula with positive *Kendall correlation*:

$$\kappa := \mathbb{E}\left[\operatorname{sign}\left(Z - \grave{Z}\right)\operatorname{sign}\left(X - \grave{X}\right)\right] > 0$$

  where $\left(\grave{Z}, \grave{X}\right)$ is an independent copy of (Z, X)
- Let (Z, X') be the associated Gaussian copula with

$$\rho = \sin\left(\pi\kappa/2\right)$$

- If (Z, X) is not 'too far' from its associated Gaussian copula:

$$\sup_{(z,x)\in(0,1)^2}\left\{\Pr\left(X' \leq x \middle| Z = z\right) - \Pr\left(X \leq x \middle| Z = z\right)\right\} \leq \nu$$

  then

$$p'_{\mathtt{success}}\left(n, m, \alpha\right) \geq p^{\mathcal{N}}_{\mathtt{success}}\left(n, m, \alpha, \rho\right) - \nu$$

Outline

Offline Controller Tuning Setup

- Controller parameter $\theta \in \Theta$
  - Candiates drawn from distribution $\mathcal{P}_\theta$
- Plant parameter $\psi \in \Psi$
  - Uncertainty with distribution $\mathcal{P}_\psi$
- Controller performance function (measurable) :

$$\overline{J}(\theta) : \Theta \to \mathbb{R}$$

- System performance function (measurable) :

$$J(\psi, \theta) : \Psi \times \Theta \to \mathbb{R}$$

- Induces a joint distribution for $\left(\overline{J}(\theta), J(\psi, \theta)\right)$

OO for Controller Tuning

- Draw n i.i.d. candidates $\theta_i \sim \mathcal{P}_\theta, \quad i = 1, \ldots, n$
- Tuned controller is best offline performer:

$$\theta^* = \underset{\theta_i \in \{\theta_1, \ldots, \theta_n\}}{\operatorname{argmin}} \; \bar{J}(\theta_i)$$

- Independently realise a plant $\psi^* \sim \mathcal{P}_\psi$ and test online:

$$J(\psi^*, \theta^*)$$

- Probability of meeting performance specification J^*

$$\Pr_{\psi^*, \theta^*}(J(\psi^*, \theta^*) \leq J^*)$$

  - Appeal to OO success probability with n samples, $m = 1$, and $\alpha = \Pr_{\psi^*, \theta_i}(J(\psi^*, \theta_i) \leq J^*)$

Sequential Learning for Controller Tuning

- In practice, copula of $\left(\bar{J}(\theta), J(\psi, \theta)\right)$ and α unknown.
- Need to estimate α, and ρ of the associated Gaussian copula
- Assumptions:
  1. $\left(\bar{J}(\theta), J(\psi, \theta)\right)$ has unknown continuous distribution with 'positive dependence' ($\kappa > 0$)
  2. Copula is not 'too far' from associated Gaussian copula (with ν)
  3. Performance specification J^* is 'feasible'
  4. Allowed to draw samples from $\mathcal{P}_\theta$, $\mathcal{P}_\psi$ and evaluate $\bar{J}$, J
- Result: sequential learning algorithm that stops after τ samples, finds a controller θ^*_τ such that

$$\Pr_{\psi^*, \theta^*_\tau}\left(J(\psi^*, \theta^*_\tau) \leq J^*\right) \geq 1 - \gamma$$

for any given $\gamma \in (\nu, 1]$.

Lower Confidence Bound for Success Probability

- Given i.i.d. sample of $\left(\bar{J}(\theta), J(\psi, \theta)\right)$
- Obtain **finite-sample** *lower confidence bounds* for α and ρ:

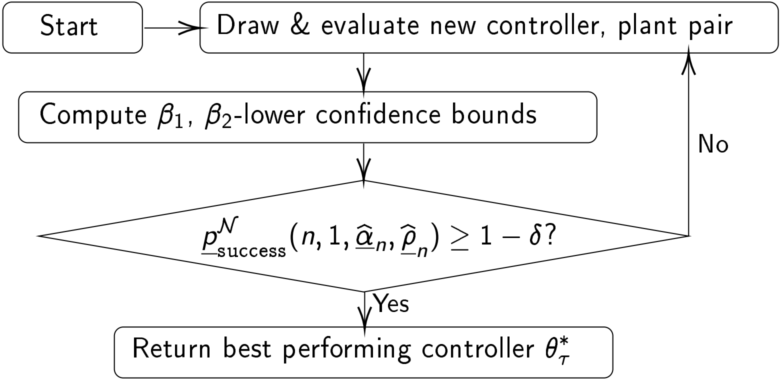
$$\Pr\left(\alpha > \underline{\widehat{\alpha}}_n\right) \geq 1 - \beta_1$$
$$\Pr\left(\rho > \underline{\widehat{\rho}}_n\right) \geq 1 - \beta_2$$

- Combine with monotonicity properties and lower bound:

$$\Pr\left(p_{\text{success}}(n, m, \alpha) \geq \underline{p}^{\mathcal{N}}_{\text{success}}\left(n, 1, \underline{\widehat{\alpha}}_n, \underline{\widehat{\rho}}_n\right) - \nu\right) \geq 1 - \beta_1 - \beta_2$$

Sequential Learning Algorithm

- Algorithm settings $\beta_1 \in (0,1]$, $\beta_2 \in (0,1]$, $\delta \in (0,1]$



- With settings $\delta + \beta_1 + \beta_2 = \gamma - \nu$, then

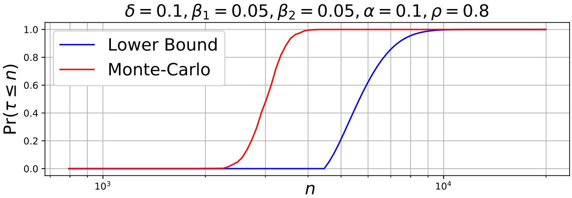
$$\Pr_{\psi^*, \theta^*_\tau}\left(J\left(\psi^*, \theta^*_\tau\right) \leq J^*\right) \geq 1 - \delta - \beta_1 - \beta_2 - \nu$$
$$= 1 - \gamma$$

Algorithm Stopping Time

- Algorithm stops after drawing a random number of samples τ
- Stopping time τ is almost surely finite:

$$\Pr(\tau \leq n) \geq 1 - O\left(e^{-\lambda n}\right)$$

  with some constant $\lambda > 0$
- Bound can be computed more precisely:



Numerical Example: MPC for Diesel Air-Path

- Step reference for output

$$y = (p_{\mathrm{im}}, y_{\mathrm{EGR}})$$

  using inputs

$$u = (u_{\mathrm{thr}}, u_{\mathrm{EGR}}, u_{\mathrm{VGT}})$$

- 4 states, with state and input constraints
- $\mathcal{P}_\psi$: Gaussian perturbations to elements of nominal model $(\overline{A}, \overline{B})$

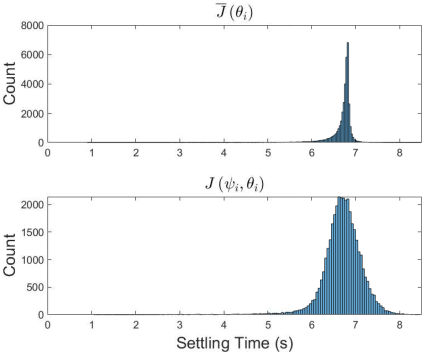
$$A = \overline{A} + S_A$$
$$B = \overline{B} + S_B$$

- $\mathcal{P}_\theta$: Spectral decomposition to generate random positive definite matrices

$$Q = W_Q D_Q W_Q^\top$$

Algorithm Results

- Performance specification: $J^* = 6.5$ seconds for settling time of $\mathrm{y}_{\mathrm{EGR}}$
- $\overline{J}$ evaluated using nominal model
- Single algorithm run with $1-\gamma = 0.7$ stopped after $\tau = 37144$ samples



Tuning Results

- Test single tuned controller θ^*_τ on 'fleet' of 10,000 independently generated plants
- All 10,000 tests met performance specification
- Similar results for multiple algorithm runs
- Possible sources of conservativeness:
  - Inherent in lower confidence bounds
  - Actual copula may be more favourable to success probabiltiy than associated Gaussian

Further Work

- Physical experiments in combination with online tuning
- Testing other classes of controllers
- Principles of choosing $\mathcal{P}_\theta$
- Lower bound that includes m
- When distribution of $\left(\overline{J}(\theta), J(\psi, \theta)\right)$ is not continuous