Learning optimal actions from experience
a neural approach

Corrado Possieri

Istituto di Analisi dei Sistemi ed Informatica “A. Ruberti”,
Consiglio Nazionale delle Ricerche, 00185 Roma, Italy

June 16, 2020
Finite-horizon optimization problem

Consider the discrete-time system

\[ x_{k+1} = f_k(x_k, u_k), \]

and the cost function

\[ J(x_0, u_0, \ldots, u_{N-1}) = \sum_{k=0}^{N} \left( \|x_k - x_k^*\|^2 + \|u_k - u_k^*\|^2 \right), \]

where \( x_k^* \) and \( u_k^* \) are reference trajectories and inputs.

Let \( U_k \) be the (convex) set of admissible inputs at time \( k \).

Optimal solution

Find a sequence \( \{u_0^*, \ldots, u_{N-1}^*\} \) that minimizes \( J \),

\[ J^*(x_0) := \min_{u_k \in U_k, k=0,\ldots,N-1} J(x_0, u_0, \ldots, u_{N-1}). \]
An example
Artificial pancreas

Artificial Pancreas

("Closed-loop" controls)

Desired glucose level

Glucose error signal = desired glucose level – actual glucose level

Blood glucose feedback

Sensed glucose level

Continuous blood glucose monitoring

Insulin Pump

Gain (insulin dose), bolus shaping (timing), adaptive filtering (learn from past experiences)

Predicted Insulin Delivery

Actual Insulin Need

Actual Endocrine Physiology

Dynamic physiology (eating, exercise, stress and illness)

Prevention of Long-term complications

Happy User!

Actual glucose metabolism
The classical solution to the problem consists in solving

$$\min \sum_{k=0}^{N} (\|x_k - x_k^*\|^2 + \|u_k - u_k^*\|^2),$$

with

$$x_{k+1} = f_k(x_k, u_k), \quad k = 0, \ldots, N - 1,$$

$$u_k \in U_k, \quad k = 0, \ldots, N - 1,$$

feed the system with $u_0$ and repeat in a rolling horizon fashion.

Main drawbacks:

Curse of modeling: availability of an accurate model;

Curse of dimensionality: computational cost.
Possible solution to the modeling issue

Use a feedforward neural network.

Universal approximation theorem

For each real-valued continuous function \( g(\cdot) \) defined on a compact, nonempty set \( \mathcal{K} \subset \mathbb{R}^n \), and for each \( \varepsilon > 0 \) there exists a feedforward neural network \( \phi(\cdot) \) such that

\[
|\phi(x) - g(x)| \leq \varepsilon, \text{forall } x \in \mathcal{K}.
\]

Use data to approximate

\[
x_0, u_0, \ldots, u_{N-1} \xrightarrow{\phi} J(x_0, u_0, \ldots, u_{N-1}).
\]
The computational issue persists

The critical issue of solving

\[
\min \phi(x_0, u_0, \ldots, u_{N-1}),
\]

with \( u_k \in U_k, \quad k = 0, \ldots, N - 1, \)

is that it is a \textbf{non-convex} optimization problem:

- extremely large computation times;
- sub-optimality of the solution;
- curse of dimensionality;
- local solutions.
DLSE\textsubscript{T} neural networks

\[
\begin{align*}
\ell^{(\alpha,\beta/T)}(x/T) & \quad d^{(\alpha,\beta,\delta)}_T(x) \\
\ell^{(\gamma,\delta/T)}(x/T) & 
\end{align*}
\]
Universal approximation result

Functions in DLSE$_T$ are$^1$

- smooth;
- in a difference of **convex** form.

Universal approximation

Given a continuous $g(\cdot)$ defined on a compact convex set $\mathcal{K} \subset \mathbb{R}^n$, for all $\varepsilon > 0$ there exist $T > 0$ and $d_T \in$ DLSE$_T$ such that

$$|d_T(x) - g(x)| \leq \varepsilon, \quad \text{forall } x \in \mathcal{K}.$$ 

We can still use data to approximate

$$x_0, u_0, \ldots, u_{N-1} \xrightarrow{d_T} J(x_0, u_0, \ldots, u_{N-1}).$$

---

DLSE_T networks are easily trainable

A DLSE_T network is a function of the parameters $\alpha, \gamma, \beta, \delta$.

We can compute the gradients

$$\nabla_{\alpha^{(i)}} d_T^{(\alpha, \gamma, \beta, \delta)}(x) = \frac{\exp(\langle \alpha^{(i)}, x/T \rangle + \beta_i/T) x}{\sum_{k=1}^{K} \exp(\langle \alpha^{(k)}, x/T \rangle + \beta_k/T)},$$

$$\nabla_{\beta_i} d_T^{(\alpha, \gamma, \beta, \delta)}(x) = \frac{\exp(\langle \alpha^{(i)}, x/T \rangle + \beta_i/T)}{\sum_{k=1}^{K} \exp(\langle \alpha^{(k)}, x/T \rangle + \beta_k/T)},$$

$$\nabla_{\gamma^{(i)}} d_T^{(\alpha, \gamma, \beta, \delta)}(x) = -\frac{\exp(\langle \gamma^{(i)}, x/T \rangle + \delta_i/T) x}{\sum_{k=1}^{K} \exp(\langle \gamma^{(k)}, x/T \rangle + \delta_k/T)},$$

$$\nabla_{\delta_i} d_T^{(\alpha, \gamma, \beta, \delta)}(x) = -\frac{\exp(\langle \gamma^{(i)}, x/T \rangle + \delta_i/T)}{\sum_{k=1}^{K} \exp(\langle \gamma^{(k)}, x/T \rangle + \delta_k/T)},$$

and use classical training algorithms to train these networks.
$\textbf{DLSE}_T$ network are easily optimizable

\textbf{Input:} functions $g_T = f_T^{(\alpha, \beta)}$ and $h_T = f_T^{(\gamma, \delta)}$ in LSE$_T$ and a convex compact set $\mathcal{K}$.

\textbf{Output:} a candidate optimal solution $\hat{x}^*$ to the problem

$$\min_{x \in \mathcal{K}} (g_T(x) - h_T(x)).$$

1: pick initial point $\chi^{(0)} \in \mathcal{K}$,
2: \textbf{for} $\kappa \in \mathbb{N}$ \textbf{do}
3: \quad let $v^{(\kappa)} = \frac{\sum_{k=1}^{K} \exp(\langle \gamma^{(k)}, \chi^{(\kappa)} \rangle / T + \delta_k / T) \gamma^{(k)}}{\sum_{k=1}^{K} \exp(\langle \gamma^{(k)}, \chi^{(\kappa)} \rangle / T + \delta_k / T)}$
4: \quad let $\chi^{(\kappa+1)} = \arg \min_{x \in \mathcal{K}} \{g_T(x) - \langle x, v^{(\kappa)} \rangle\}$,
5: \quad \textbf{if} $\frac{\|\chi^{(\kappa+1)} - \chi^{(\kappa)}\|}{1 + \|\chi^{(\kappa)}\|}$ is smaller than a tolerance \textbf{then}
6: \quad \textbf{return} $\hat{x}^* = \chi^{(\kappa+1)}$. 
Letting $\hat{u}$ be the solution to

$$\min d_T(x_0, u_0, \ldots, u_{N-1}),$$
with $u_k \in U_k, \ k = 0, \ldots, N - 1,$

and letting $u$ be the solution to

$$\min \sum_{k=0}^{N} (\|x_k - x_k^*\|^2 + \|u_k - u_k^*\|^2),$$
with $x_{k+1} = f_k(x_k, u_k), \ k = 0, \ldots, N - 1,$
$u_k \in U_k, \ k = 0, \ldots, N - 1,$

we have that

$$|\hat{u} - u| < \epsilon, \ for \ all \ x_0 \in \mathcal{X}.$$
Lithium ion distribution

Consider the model for the distribution of Lithium ions in humans.

The objective is to let the concentration levels $C$ satisfy

\[ 0.4 \text{ nmol/L} \leq C_{\text{plasma}} \leq 0.6 \text{ nmol/L}, \]
\[ 0.6 \text{ nmol/L} \leq C_{\text{RBC}} \leq 0.9 \text{ nmol/L}, \]
\[ 0.5 \text{ nmol/L} \leq C_{\text{muscle}} \leq 0.8 \text{ nmol/L}. \]
Test of the $\text{LSE}_T$ approach

Table: Average costs

<table>
<thead>
<tr>
<th>Epoch</th>
<th>LSE networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.9925</td>
</tr>
<tr>
<td>2</td>
<td>2.4126</td>
</tr>
<tr>
<td>3</td>
<td>2.2527</td>
</tr>
<tr>
<td>4</td>
<td>1.9026</td>
</tr>
<tr>
<td>5</td>
<td>1.6736</td>
</tr>
<tr>
<td>6</td>
<td>1.6430</td>
</tr>
</tbody>
</table>
Do you have problems with data, no model, and an intrinsic optimization to be performed?

Mail to: corrado.possieri@iasi.cnr.it.