Koopman transfer learning via perturbation

Preprint · January 2024

DOI: 10.13140/RG.2.2.21890.66248/2

Gary Nan Tie
Mu Risk LLC

Koopman transfer learning via perturbation

Gary Nan Tie
Jan 16, 2024

Abstract
We use a RKHS Representer Theorem to construct a vector-valued Koopman operator, and then give an analytic representation of a Koopman embedding, that has a novel invertible perturbation used to predict future states of nonlinear dynamical systems.

'More or less, less is more!'

Keywords: nonlinear dynamical systems, Koopman operators, Koopman embeddings future state prediction, reproducing kernel Hilbert space, Representer Theorem, rank-one perturbations, data driven discovery, machine learning

Preliminaries

Koopman operators:

Let $f : \mathcal{X} \to \mathcal{X}$, $\quad \mathcal{X} \subseteq \mathbb{R}^{n_x}$,

and $\quad x_{k+1} = f(x_k)$, $\quad k \in \mathbb{Z}_+$

Let $\mathcal{H} = \mathcal{H}(\mathbb{k}) \subseteq \mathrm{Fun}(\mathcal{X}, \mathbb{R})$ be a RKHS

with reproducing kernel $\mathbb{k} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$

Assume $g \circ f \in \mathcal{H}$, $\quad \forall g \in \mathcal{H}$.

Linear map $K : \mathcal{H} \to \mathcal{H}$ is said to be a

Koopman operator if $\quad Kg = g \circ f$, $\quad \forall g \in \mathcal{H}$.

Let X be a finite or countable set,

Hilbert space $\ell^2(X) \triangleq \{ f : X \to \mathbb{C} \mid \sum_{x \in X} |f(x)|^2 < +\infty \}$

where $\langle f, g \rangle \triangleq \sum_{x \in X} f(x) \overline{g(x)}$

Let $e_y(x) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$, then $\{e_y\}_{y \in X}$ is an orthonormal basis for $\ell^2(X)$ and $\langle f, e_y \rangle = f(y)$.

$\Bbbk(x, y) \triangleq \langle e_x, e_y \rangle$ is reproducing kernel for $\ell^2(X)$.

$\mathcal{H}(\Bbbk) = \ell^2(X)$ is a reproducing kernel Hilbert space.

Vector-valued RKHS:

Let $\mathcal{C}$ be a Hilbert space and $\mathrm{Fun}(X, \mathcal{C})$ be the vector space of $\mathcal{C}$-valued functions on X under pointwise sum and scalar multiplication. A subspace $\mathcal{H} \subseteq \mathrm{Fun}(X, \mathcal{C})$ is called a <u>$\mathcal{C}$-valued RKHS on X</u> provided $\mathcal{H}$ is a Hilbert space and $\forall y \in X$, the linear evaluation map

$E_y : \mathcal{H} \to \mathcal{C}$, $\quad E_y(f) = f(y)$ is bounded.

For example, $\mathcal{C} = \mathbb{C}^n$, $\quad \langle v, w \rangle = \sum_{i=1}^{n} v_i \overline{w_i}$

and $\mathcal{H}(k) \subseteq \mathrm{Fun}(X, \mathbb{C})$ a scalar-valued RKHS on X.

Let $\mathcal{H}^n = \bigoplus_n \mathcal{H}$ with $\langle f, g \rangle_{\mathcal{H}^n} = \sum_{i=1}^{n} \langle f_i, g_i \rangle_{\mathcal{H}}$

Define
$$\begin{aligned} L : \mathcal{H}^n &\hookrightarrow \mathrm{Fun}(X, \mathbb{C}^n) \\ f &\longmapsto x \mapsto (f_1(x), \dots, f_n(x)) \end{aligned}$$

Identify $\mathcal{H}^n$ with $L(\mathcal{H}^n)$ then $\|E_y\| = \|k_y\|$.

<u>So $\mathcal{H}^n$ is a $\mathbb{C}^n$-valued RKHS.</u>

Representer Theorem for Koopman operators:

RKHS $\mathcal{H} = \mathcal{H}(\Bbbk) \subseteq \mathrm{Fun}(\mathcal{X}, \mathbb{R})$, $\mathcal{X} \subseteq \mathbb{R}^{n_x}$

with reproducing kernel $\Bbbk : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.

Learning problem

$$\hat{K} \triangleq \min_{K \in \mathcal{L}(\mathcal{H})} \sum_{k=1}^{n_s} \sum_{l=1}^{n_s} \left[y_{kl} - (K g_l)(x_{k-1}) \right]^2 + \lambda \| K \|^2$$

where $y_{kl} \triangleq g_l(x_k)$, $k = 0, \ldots, n_s$, $l = 1, \ldots, n_s$ and $\lambda > 0$.

Representer Theorem [Khosravi]

The learning problem has a unique solution.

Moreover, there exist $v_1, \ldots, v_{n_s} \in \mathcal{H}$ such that

$$\hat{K} = \sum_{k=1}^{n_s} \sum_{l=1}^{n_s} \alpha_{kl} \, v_k \otimes g_l \,, \quad \text{where } A = [\alpha_{kl}] \in \mathbb{R}^{n_s \times n_s}$$

solves: $$\min_{A \in \mathbb{R}^{n_s \times n_s}} \| VAG - Y \|_{Frob}^2 + \lambda \| V^{\frac{1}{2}} A G^{\frac{1}{2}} \|^2$$

with V and G respectively the Gramian matrix of

$\{v_1, \ldots, v_{n_s}\}$ and $\{g_1, \ldots, g_{n_s}\}$ and $Y = [y_{kl}] \in \mathbb{R}^{n_s \times n_s}$.

In fact, $v_k = \Bbbk(x_{k-1}, \cdot)$ for $k \in [n_s]$.

<u>Lemma</u> [Ionascu, Integral Eqns & Op Thy, Dec 2001, 39, 421-440]

Let $\mathcal{L}(\mathcal{H})$ be the algebra of bounded linear operators on Hilbert space $\mathcal{H}$. Let $A \in \mathcal{L}(\mathcal{H})$ be invertible and

$S = A + u \otimes v$ be a rank-one perturbation.

S is invertible iff $\langle A^{-1}u, v \rangle \neq -1$.

Moreover, $S^{-1} = A^{-1} - \dfrac{1}{\langle A^{-1}u, v\rangle + 1} \left(A^{-1}u \otimes (A^*)^{-1} v \right)$

$\not\!\!-$

$$\hat{K} = \sum_{k=1}^{n_s} \Big(\sum_{l=1}^{n_s} \alpha_{kl} \, v_k \otimes g_l \Big) = \sum_{j=1}^{n_s \times n_s} \varphi_j \otimes g_j$$

after reindexing in lex order with $\varphi = \alpha v$.

Let $A_0 \triangleq \lambda I$, $\lambda > 0$, and for $j = 1, \ldots, n_s \times n_s$

$$A_j \triangleq A_{j-1} + \varphi_j \otimes \mathfrak{g}_j, \text{ where } \mathfrak{g}_j = \begin{cases} g_j & \text{if } \langle A_{j-1}^{-1} \varphi_j, g_j \rangle \neq -1 \\ (1+\varepsilon) g_j, & \varepsilon > 0, \text{ otherwise} \end{cases}$$

Then by the lemma,

$$K^\varphi \triangleq \lambda I + \sum_{j=1}^{n_s \times n_s} \varphi_j \otimes \mathfrak{g}_j \text{ is an invertible perturbation of } \hat{K}.$$

<u>Prop</u> $\quad \| K^{\circ} - \hat{K} \| \leqslant \lambda + \varepsilon \sum_{j \in J} \| \varphi_j \| \, \| g_j \|$

where $\quad J \subseteq [n_s \times n_s]$

<u>Pf</u>

$$\mathfrak{g}_j \triangleq \begin{cases} g_j & \text{if } \langle A_{j-1}^{-1} \varphi_j , g_j \rangle \neq -1 \\ (1+\varepsilon) g_j , & \varepsilon > 0 , \text{ otherwise and say } j \in J \end{cases}$$

$$\| K^{\circ} - \hat{K} \|$$

$$= \Big\| \lambda I + \sum_{j=1}^{n_s \times n_s} (\varphi_j \otimes \mathfrak{g}_j - \varphi_j \otimes g_j) \Big\|$$

$$\leqslant \lambda + \sum_{j \in J} \| \varphi_j \otimes (1+\varepsilon) g_j - \varphi_j \otimes g_j \|$$

$$= \lambda + \sum_{j \in J} \varepsilon \| \varphi_j \otimes g_j \|$$

$$= \lambda + \varepsilon \sum_{j \in J} \| \varphi_j \| \, \| g_j \|$$

where $\lambda > 0$ and $\varepsilon > 0$.

$\square$

So <u>K° is an invertible perturbation of $\hat{K}$</u> that with respect to operator norm can approximate $\hat{K}$ arbitrarily closely.

Vector-valued Koopman operator theory:

Unknown dynamics $f: \mathcal{X} \to \mathcal{X}$, finite $\mathcal{X} \subset \mathbb{R}^{n_x}$

$$x_{k+1} = f(x_k) \quad , \quad k = 0, 1, \ldots, n_s$$

$$\mathcal{H} \triangleq \ell^2(\mathcal{X}) \triangleq \left\{ h: \mathcal{X} \to \mathbb{R} \;\middle|\; \sum_{x \in \mathcal{X}} |h(x)|^2 < \infty \right\}$$

is a RKHS with inner product $\langle f, g \rangle \triangleq \sum_{x \in \mathcal{X}} f(x) g(x)$

and reproducing kernel $\mathbb{k}(x, y) \triangleq \langle e_x, e_y \rangle$.

Note that
$$\begin{aligned} \pi_i : \mathcal{X} &\longrightarrow \mathbb{R} \quad \in \mathcal{H}. \\ (y_1, \ldots, y_{n_x}) &\longmapsto y_i \end{aligned}$$

Observable maps $g_l \triangleq \mathbb{k}(p_l, \cdot)$, $p_l \in \mathcal{X}$, $l \in [n_g]$

with measurements $y_{kl} \triangleq g_l(x_k)$.

Define Koopman operators $\hat{K}_i : \mathcal{H} \longrightarrow \mathcal{H}$, $i \in [n_x]$

by $$\hat{K}_i \triangleq \min_{K \in \mathcal{L}(\mathcal{H})} \sum_{k=1}^{n_s} \sum_{l=0}^{n_g} \left[g_l(x_k) - (K g_l)(x_{k-1}) \right]^2 + \lambda \|K\|^2, \quad \lambda > 0$$

where: $g_0 \triangleq \pi_i$ and $g_l = k(p_l, \cdot)$, $l \in [n_g]$.

By the Representer Theorem:

$$\hat{K}_i = \sum_{k=1}^{n_s} \sum_{l=0}^{n_g} \alpha_{kli} \, v_k \otimes g_l, \quad v_k = k(x_{k-1}, \cdot) \; k \in [n_s], \; g_0 = \pi_i$$

So $g_l(x_k) \approx (\hat{K}_i g_l)(x_{k-1})$, $l \in [n_g]$ and $\pi_i(x_k) \approx (\hat{K}_i \pi_i)(x_{k-1})$.

<u>Vector-valued Koopman operator</u> $\hat{\mathbb{K}} : \mathcal{H}^{n_x} \longrightarrow \mathcal{H}^{n_x}$

$$\hat{\mathbb{K}} \triangleq \bigoplus_{i \in [n_x]} \hat{K}_i : \bigoplus_{n_x} \mathcal{H} \longrightarrow \bigoplus_{n_x} \mathcal{H}$$

<u>Koopman embedding</u> $\varphi : \mathcal{X} \longrightarrow \mathbb{R}^{n_x}$, $\mathcal{X} \subset \mathbb{R}^{n_x}$

$\varphi \triangleq [\varphi_i] \in \mathcal{H}^{n_x} \subseteq \mathrm{Fun}(\mathcal{X}, \mathbb{R}^{n_x})$, finite $\mathcal{X} \subset \mathbb{R}^{n_x}$

where $\varphi_i \triangleq \hat{K}_i \pi_i \in \mathcal{H}$, $i \in [n_x]$.

Note that $\hat{K}_i$ solves the learning problem with

$v_k = k(x_{k-1}, \cdot)$, $k \in [n_s]$ and $g_l = k(p_l, \cdot)$, $l \in [n_g]$

under the <u>constraint</u> $\pi_i(x_k) = (K \pi_i)(x_{k-1})$, $k \in [n_s]$.

Koopman embedding φ is not necessarily invertible,

however perturbation proxy $\phi \triangleq [\, K^{\circ}_i \, \pi_i \,]$ is invertible,

with inverse $\psi \triangleq [\, K^{\circ -1}_i \, \pi_i \,]$.

$$\begin{aligned}
\psi \, \phi(x) &= [\, \dots, \psi_i[\, \phi_1(x), \dots, \phi_{n_x}(x) \,], \dots \,] \\
&= [\, \dots, K^{\circ -1}_i \, \phi_i(x), \dots \,] \\
&= [\, \dots, K^{\circ -1}_i \, K^{\circ}_i \, \pi_i(x), \dots \,] \\
&= [\, \dots, x_i, \dots \,] = x \in \mathcal{X}
\end{aligned}$$

The Representer Theorem constraint for $\hat{K}_i$ implies

$\hat{K}_i \, \varphi_i(x_{k-1}) \approx \varphi_i(x_k)$ ie the linearity of $\hat{K}$.

Define future state $\hat{x}_{k+1} \triangleq \phi^{-1}(\hat{K} \varphi(x_k))$, $k = n_s, n_s+1, \dots$

$$\begin{tikzcd}
x_k \arrow[r, "f"] \arrow[d, "\varphi"'] & x_{k+1} \arrow[d, "\varphi \approx \phi"'] \\
\varphi(x_k) \arrow[r, "\hat{K}"'] & \varphi(x_{k+1}) \arrow[u, bend right, "\psi = \phi^{-1}"']
\end{tikzcd}$$

<u>Summary:</u>

Koopman operators $\hat{K}_i : \mathcal{H} \to \mathcal{H}$, $\hat{K} = \bigoplus_{i \in [n_x]} \hat{K}_i : \mathcal{H}^{n_x} \to \mathcal{H}^{n_x}$,

Koopman embedding $\varphi : \mathcal{X} \to \mathbb{R}^{n_x}$,

and invertible perturbation $\psi : \mathcal{X} \to \mathbb{R}^{n_x}$

have been designed so that:

i) $g_i(x_k) \approx (\hat{K}_i g_i)(x_{k-1})$ Learning Problem

ii) $\varphi_i(x_k) \approx (\hat{K}_i \varphi_i)(x_{k-1})$ Linearity

iii) $\psi^{-1}(\psi(x_k)) = x_k$ Autoencode

iv) $x_{k+1} \approx \psi^{-1}(\hat{K}\,\psi(x_k))$ Prediction

v) $\psi \approx \varphi$ Perturbation

Synopsis: Given data:

a) trajectory $x_0, x_1, \ldots, x_{n_s} \in \mathbb{R}^{n_x}$

b) observable maps $g_1, \ldots, g_{n_g} \in \mathcal{H} \triangleq \ell^2(\mathcal{X})$

where $g_l \triangleq k(p_l, \cdot)$, $p_l \in$ finite $\mathcal{X} \subset \mathbb{R}^{n_x}$

c) measurements $y_{kl} \triangleq g_l(x_k)$, $k = 0, \ldots, n_s$, $l = 1, \ldots, n_g$.

1. Determine vector-valued Koopman operator:

$$\hat{K} \triangleq \bigoplus_{i \in [n_x]} \hat{K}_i : \mathcal{H}^{n_x} \longrightarrow \mathcal{H}^{n_x}$$

by the constrained Representer Theorems for $\hat{K}_i : \mathcal{H} \to \mathcal{H}$, $(g_0 = \pi_i)$.

2. Determine Koopman embedding:

$$\varphi = [\varphi_i] : \mathcal{X} \longrightarrow \mathbb{R}^{n_x} \quad \text{where} \quad \varphi_i \triangleq \hat{K}_i \pi_i$$

3. Calculate invertible perturbation $\tilde{\varphi} \triangleq [\tilde{K}_i \pi_i]$ to proxy φ.

4. Predict future states $\hat{x}_{k+1} \triangleq \tilde{\varphi}^{-1}(\hat{K} \varphi(x_k))$,

$k = n_s, n_s + 1, \ldots$

References

Khosravi, M.
'Representer Theorem for Learning Koopman Operators'
IEEE Transactions on Automatic Control,
Vol. 68, No. 5, May 2023

V. Paulsen, M. Raghupathi,
"An Introduction to the Theory of Reproducing Kernel
Hilbert Spaces",
Cambridge University Press, 2016.
ISBN 978-1-107-10409-9