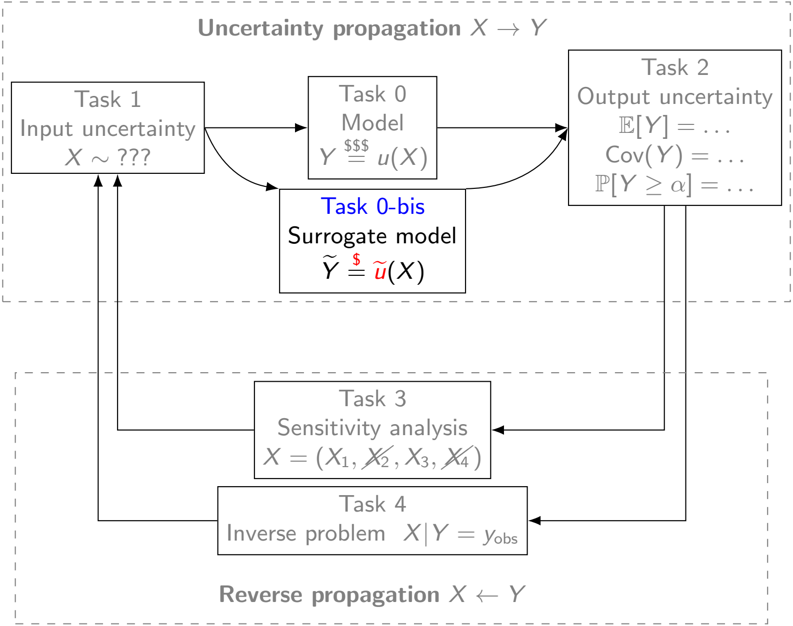
Poincaré inequalities for dimension reduction and efficient sampling

Part I: The curse of dimensionality

Olivier Zahm
olivier.zahm@inria.fr

INRIA – Université Grenoble-Alpes
https://team.inria.fr/airsea/en/

ETICS
October 2025


Uncertainty propagation $X \to Y$
Task 1
Input uncertainty
$X \sim ???$
Task 0
Model
$Y \overset{\$\$\$}{=} u(X)$
Task 0-bis
Surrogate model
$\widetilde{Y} \overset{\$}{=} \widetilde{u}(X)$
Task 2
Output uncertainty
$\mathbb{E}[Y] = \dots$
$\mathrm{Cov}(Y) = \dots$
$\mathbb{P}[Y \geq \alpha] = \dots$
Task 3
Sensitivity analysis
$X = (X_1, \cancel{X_2}, X_3, \cancel{X_4})$
Task 4
Inverse problem $X|Y = y_{\mathrm{obs}}$
Reverse propagation $X \leftarrow Y$

The **curse of dimensionality** in function approximation

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \dots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \leq \varepsilon \|u\|_\infty$$

How many point evaluations of u do I need?

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \ldots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \leq \varepsilon \|u\|_\infty$$

  How many point evaluations of u do I need?
- **A:** Well, if u is linear, then $N = d + 1$ evaluations should be enough...

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \ldots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \leq \varepsilon \|u\|_\infty$$

  How many point evaluations of u do I need?
- **A:** Well, if u is linear, then $N = d + 1$ evaluations should be enough...
- **Q:** Okay... what if u is *just* extremely regular, say,

$$\sup_{\alpha \in \mathbb{N}^d} \left\| \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_\infty < \infty$$

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \ldots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \leq \varepsilon \|u\|_\infty$$

  How many point evaluations of u do I need?
- **A:** Well, if u is linear, then $N = d + 1$ evaluations should be enough...
- **Q:** Okay... what if u is *just* extremely regular, say,

$$\sup_{\alpha \in \mathbb{N}^d} \left\| \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_\infty < \infty$$

- **A:** Sorry: for any approximation methods you can ever think of, there exists such a u which would require at least

$$\boxed{\boxed{N \geq 2^{\lfloor d/2 \rfloor}}}$$

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \ldots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \le \varepsilon \|u\|_\infty$$

  How many point evaluations of u do I need?
- **A:** Well, if u is linear, then $N = d + 1$ evaluations should be enough...
- **Q:** Okay... what if u is *just* extremely regular, say,

$$\sup_{\alpha \in \mathbb{N}^d} \left\| \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_\infty < \infty$$

- **A:** Sorry: for any approximation methods you can ever think of, there exists such a u which would require at least

$$\boxed{\boxed{N \ge 2^{\lfloor d/2 \rfloor}}}$$

- **Q:** Come on! I remember that with polynomial interpolation, I can reach

$$\|u - \widetilde{u}_{\text{interpolation}}\|_\infty \le \mathcal{O}(\rho^N)$$

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

The **curse of dimensionality** in function approximation

- **Q:** I want to construct an approximation $\widetilde{u}$ to $u(x_1, \ldots, x_d)$ such that

$$\|u - \widetilde{u}\|_\infty \leq \varepsilon \|u\|_\infty$$

  How many point evaluations of u do I need?
- **A:** Well, if u is linear, then $N = d + 1$ evaluations should be enough...
- **Q:** Okay... what if u is *just* extremely regular, say,

$$\sup_{\alpha \in \mathbb{N}^d} \left\| \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_\infty < \infty$$

- **A:** Sorry: for any approximation methods you can ever think of, there exists such a u which would require at least

$$\boxed{\boxed{N \geq 2^{\lfloor d/2 \rfloor}}}$$

- **Q:** Come on! I remember that with polynomial interpolation, I can reach

$$\|u - \widetilde{u}_{\text{interpolation}}\|_\infty \leq \mathcal{O}(\rho^N)$$

- **A:** Sure. But the constant hidden in $\mathcal{O}$ depends exponentially on d. You'll need at least $N \geq 2^{\lfloor d/2 \rfloor}$ to make sure you reach the asymptotic regime.

[Novak and Woźniakowski 2009] **Approximation of infinitely differentiable multivariate functions is intractable**, Journal of Complexity

Take-away message: **regularity is not enough!**

High-dimensional function approximation necessitates leveraging any **underlying low-dimensional structure** present in u, such as:

- **Sparse structure**

$$u(x) \approx \sum_{\alpha \in \Lambda_N} u_\alpha \Phi_\alpha(x), \qquad \begin{cases} \Phi_\alpha(x): & \text{given basis} \\ \#\Lambda_N = N \end{cases}$$

- **Low-rank structure**

$$u(x) \approx \sum_{i=1}^{r} u_1^i(x_1) \dots u_d^i(x_d)$$

- **Compositional structure**

$$u(x) \approx f \circ g(x), \quad \begin{cases} g: \mathbb{R}^d \to \mathbb{R}^m \\ f: \mathbb{R}^m \to \mathbb{R} \\ m: & \text{latent dimension} \end{cases}$$

- ...

A prototypical example: **parametrized elliptic PDE**

Find $u(x) \in H^1(\Omega)$ solution to

$$-\mathrm{div}(a(x)\nabla u(x)) = f \qquad \text{in } \Omega \subset \mathbb{R}^2 \text{ or } \mathbb{R}^3$$

where

$$a(x,s) = a_0(s) + \sum_{i=1}^{d=\infty} x_i\, a_i(s), \qquad \begin{cases} x_1, x_2, \ldots \in [-1,1] \\ s \in \Omega \end{cases}$$

A prototypical example: **parametrized elliptic PDE**

Find $u(x) \in H^1(\Omega)$ solution to

$$-\mathrm{div}(a(x)\nabla u(x)) = f \qquad \text{in } \Omega \subset \mathbb{R}^2 \text{ or } \mathbb{R}^3$$

where

$$a(x,s) = a_0(s) + \sum_{i=1}^{d=\infty} x_i\, a_i(s), \qquad \begin{cases} x_1, x_2, \ldots \in [-1,1] \\ s \in \Omega \end{cases}$$

Assume

$$(\|a_i\|_\infty)_{i\geq 1} \in \ell^p \qquad \text{for some } p < 1$$

then there exists

$$\widetilde{u}(x,s) = \sum_{i=1}^{n} \varphi_i(x) v_i(s), \qquad \begin{cases} v_i \in H^1(\Omega) \\ \varphi_i \in L^\infty([-1,1]) \end{cases}$$

such that

$$\boxed{\sup_{x\in[-1,1]^{\mathbb{N}}} \|u(x) - \widetilde{u}(x)\|_{H^1(\Omega)} \leq C n^{-s}, \quad \text{where } s = p^{-1} - 1 > 0}$$

for some constant C **independent on the dimension** (no curse of dimensionality!)

[Cohen DeVore 2015: Approximation of high-dimensional parametric PDEs, Acta Numerica]

Near-optimal approximations can be obtained using

- **sparse polynomial** expansions:

$$\widetilde{u}(x,s) = \sum_{\substack{\alpha\in\Lambda_n \\ \#\Lambda_n=n}} \varphi_\alpha(x) v_\alpha(s) \qquad \begin{cases} \varphi_\alpha(x) : \text{given multivariate polynomials} \\ \Lambda_n : \textbf{Greedy algorithm } \Lambda_{n+1} = \Lambda_n \cup \{\alpha^*_{n+1}\} \\ v_\alpha(s) : \text{least-squares, interpolation, ...} \end{cases}$$

- the **Reduced Basis** method:

$$\widetilde{u}(x,s) = \sum_{i=1}^{n} \varphi_i(x) u(x_i, s) \qquad \begin{cases} x_1,\dots,x_n : \textbf{Greedy algorithm } n \leftarrow n+1 \\ \varphi_i(x) : \text{Galerkin projection of } u(x) \text{ on} \\ \qquad \text{span}(u(x_1),\dots,u(x_n)) \end{cases}$$

[Rozza, Huynh and Patera 2008] *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive PDE* Arch. Comput. Methods Eng.
[Blatman and Sudret 2011] Adaptive sparse polynomial chaos expansion based on least angle regression
[Chkifa, Cohen and Schwab 2015] *Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs* Journal de Mathématiques Pures et Appliquées
[...............................]

Illustration on Kernel ridge regression

The survival kit for **Reproducing Kernel Hilbert Space**

$$\mathcal{H} = \left\{ f : \mathbb{R}^d \to \mathbb{R},\ \|f\|^2_{\mathcal{H}} = \langle f, f \rangle^2_{\mathcal{H}} < \infty \right\}$$

$$\text{s.t.}\quad \mathcal{H} \ni f \mapsto f(x) \in \mathbb{R} \quad \textbf{continuous}\ \forall x$$

The survival kit for **Reproducing Kernel Hilbert Space**

$$\mathcal{H} = \left\{ f : \mathbb{R}^d \to \mathbb{R}, \ \|f\|^2_{\mathcal{H}} = \langle f, f \rangle^2_{\mathcal{H}} < \infty \right\}$$

$$\text{s.t.} \quad \mathcal{H} \ni f \mapsto f(x) \in \mathbb{R} \quad \textbf{continuous}\ \forall x$$

Riesz representation theorem: there exists a feature map $\varphi : x \mapsto \varphi(x) \in \mathcal{H}$ such that

$$f(x) = \langle f, \varphi(x) \rangle_{\mathcal{H}}$$

The survival kit for **Reproducing Kernel Hilbert Space**

$$\mathcal{H} = \left\{ f : \mathbb{R}^d \to \mathbb{R},\ \|f\|^2_{\mathcal{H}} = \langle f, f \rangle^2_{\mathcal{H}} < \infty \right\}$$

$$\text{s.t.}\quad \mathcal{H} \ni f \mapsto f(x) \in \mathbb{R} \quad \textbf{continuous } \forall x$$

Riesz representation theorem: there exists a feature map $\varphi : x \mapsto \varphi(x) \in \mathcal{H}$ such that

$$f(x) = \langle f, \varphi(x) \rangle_{\mathcal{H}}$$

Define the kernel $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ by

$$\kappa(x, x') := \langle \varphi(x), \varphi(x') \rangle_{\mathcal{H}}$$

which is a symmetric positive definite (SPD) function, *i.e.* $\forall x_1, x_2, \ldots \subset \mathbb{R}^d$:

$$[\kappa(x_i, x_j)]_{ij} \succeq 0$$

The survival kit for **Reproducing Kernel Hilbert Space**

$$\mathcal{H} = \left\{ f : \mathbb{R}^d \to \mathbb{R},\ \|f\|_{\mathcal{H}}^2 = \langle f, f \rangle_{\mathcal{H}}^2 < \infty \right\}$$

$$\text{s.t.}\quad \mathcal{H} \ni f \mapsto f(x) \in \mathbb{R} \quad \textbf{continuous}\ \forall x$$

Riesz representation theorem: there exists a feature map $\varphi : x \mapsto \varphi(x) \in \mathcal{H}$ such that

$$f(x) = \langle f, \varphi(x) \rangle_{\mathcal{H}}$$

Define the kernel $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ by

$$\kappa(x, x') := \langle \varphi(x), \varphi(x') \rangle_{\mathcal{H}}$$

which is a symmetric positive definite (SPD) function, *i.e.* $\forall x_1, x_2, \ldots \subset \mathbb{R}^d$:

$$[\kappa(x_i, x_j)]_{ij} \succeq 0$$

Since $\kappa(x, \cdot) = \langle \varphi(x), \varphi(\cdot) \rangle_{\mathcal{H}} = \varphi(x)(\cdot)$, we deduce the **reproducing property**

$$f(x) = \langle f, \kappa(x, \cdot) \rangle_{\mathcal{H}}$$

so that κ **entirely characterizes** $\mathcal{H}$:

$$\mathcal{H} = \left\{ f(x) = \sum_{i=1}^{\infty} \alpha_i \kappa(x, x_i) \;\middle|\; \begin{array}{l} x_i \in \mathbb{R}^d \text{ and } \alpha_1, \alpha_2, \ldots \in \mathbb{R} \text{ s.t.} \\ \|f\|_{\mathcal{H}}^2 := \sum_{i,j=1}^{\infty} \alpha_i \alpha_j\ \kappa(x_i, x_j) < \infty \end{array} \right\}$$

Kernel ridge regression: $\min \mathbb{E}[(Y - f(X))^2], \ f \in \mathcal{H} \subset L^2_{\pi_X}$

Given i.i.d. samples $\{(X^{(i)}, Y^{(i)})\}_{i=1}^N$ with $Y^{(i)} = u(X^{(i)})$, consider the regularized[1] problem

$$\min_{f \in \mathcal{H}} \ \sum_{i=1}^{d} (Y^{(i)} - f(X^{(i)}))^2 + \lambda \|f\|_{\mathcal{H}}^2$$

[1] "Ridge" or "Tikhonov" regularization: λ controls the complexity of the solution

Kernel ridge regression: $\min \mathbb{E}[(Y - f(X))^2]$, $f \in \mathcal{H} \subset L^2_{\pi_X}$

Given i.i.d. samples $\{(X^{(i)}, Y^{(i)})\}_{i=1}^N$ with $Y^{(i)} = u(X^{(i)})$, consider the regularized[1] problem

$$\min_{f \in \mathcal{H}} \sum_{i=1}^{d} (Y^{(i)} - f(X^{(i)}))^2 + \lambda \|f\|^2_{\mathcal{H}}$$

By the representer theorem, the solution necessarily writes

$$f^N(x) = \sum_{i=1}^{N} \alpha_i \; \kappa(x, X^{(i)}), \qquad \alpha \in \mathbb{R}^N$$

and the coefficients $\alpha = (\alpha_1, \dots, \alpha_N)$ are obtained via least-square regression

$$\alpha = (K + \lambda I)^{-1} Y_N \quad \text{where} \begin{cases} K = [\kappa(X^{(i)}, X^{(j)})]_{i,j} \\ Y_N = (Y^{(1)}, \dots, Y^{(N)}) \end{cases}$$

[1] "Ridge" or "Tikhonov" regularization: λ controls the complexity of the solution

Choosing the kernel

- Gaussian Radial Basis Function (infinitely smooth $\mathcal{H} \subset \mathcal{C}^\infty$)

$$\kappa(x, x') = \exp\left(-\frac{1}{2\ell^2}\|x - x'\|^2\right)$$

classically with $\ell = \mathrm{med}(\|X^{(i)} - X^{(j)}\|)$.

Choosing the kernel

- Gaussian Radial Basis Function (infinitely smooth $\mathcal{H} \subset \mathcal{C}^\infty$)

$$\kappa(x, x') = \exp\left(-\frac{1}{2\ell^2}\|x - x'\|^2\right)$$

classically with $\ell = \mathrm{med}(\|X^{(i)} - X^{(j)}\|)$.

- Matérn kernels $\kappa_\nu(x, x')$ with regularity parameter $\nu > 0$

(non-differentiable) $\quad \kappa_{1/2}(x, x') = \exp\left(-\frac{\|x - x'\|}{\ell}\right)$

(once differentiable) $\quad \kappa_{3/2}(x, x') = \left(1 + \frac{\sqrt{3}\|x - x'\|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\|x - x'\|}{\ell}\right)$

(twice differentiable) $\quad \kappa_{5/2}(x, x') = \ldots$

Choosing the kernel

- Gaussian Radial Basis Function (infinitely smooth $\mathcal{H} \subset \mathcal{C}^\infty$)

$$\kappa(x, x') = \exp\left(-\frac{1}{2\ell^2}\|x - x'\|^2\right)$$

  classically with $\ell = \mathrm{med}(\|X^{(i)} - X^{(j)}\|)$.
- Matérn kernels $\kappa_\nu(x, x')$ with regularity parameter $\nu > 0$

$$\text{(non-differentiable)} \quad \kappa_{1/2}(x, x') = \exp\left(-\frac{\|x - x'\|}{\ell}\right)$$

$$\text{(once differentiable)} \quad \kappa_{3/2}(x, x') = \left(1 + \frac{\sqrt{3}\|x - x'\|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\|x - x'\|}{\ell}\right)$$

$$\text{(twice differentiable)} \quad \kappa_{5/2}(x, x') = \dots$$

- Take your favorite basis function $\Phi(x) = (\Phi_1(x), \dots, \Phi_K(x))$, *e.g.* polynomial, wavelet, Fourier, spline,...

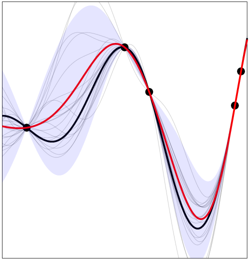
$$\kappa(x, x') = \Phi(x)^\top \Phi(x')$$

  - $N < K$ overfitting, unless $\lambda > 0$
  - $N = K$ interpolation ($\lambda = 0$)
  - $N > K$ standard least-square ($\lambda = 0$)

Link with **Gaussian Process** (*i.e.* **Kriging**)

KRR	GP
$u \in \mathcal{H}$	u is a realization of $Z \sim \mathcal{N}(0, \kappa)$
κ = kernel	$\kappa(x, x') = \mathrm{Cov}(Z(x), Z(x'))$
f^N = LS problem	$f^N(x) = \mathbb{E}[Z(X) \vert Z(X^{(i)}) = Y^{(i)}]$
λ = regulaization	λ = nugget

The **little extra** with GP is $\boxed{Z(x) \vert Z(X^{(i)}) = Y^{(i)}}$: posterior covariance, posterior samples...



Accounting for **anisotropy** in the kernel

Consider the modified Gaussian RBF

$$\kappa_B(x, x') = \exp\left(-\frac{1}{2\ell^2}\|B^\top(x - x')\|^2\right)$$

for some matrix $B \in \mathbb{R}^{d\times m}$ to determine...

$$\boxed{f \in \mathcal{H}_{\kappa_B} \qquad \Rightarrow \qquad f(x) = f(B^\top x)}$$

Accounting for **anisotropy** in the kernel

Consider the modified Gaussian RBF

$$\kappa_B(x, x') = \exp\left(-\frac{1}{2\ell^2}\|B^\top(x - x')\|^2\right)$$

for some matrix $B \in \mathbb{R}^{d \times m}$ to determine...

$$\boxed{f \in \mathcal{H}_{\kappa_B} \qquad \Rightarrow \qquad f(x) = f(B^\top x)}$$

- **[Fukumizu 2004]**: optimize over $B \in \mathbb{R}^{d \times m}$

$$\text{“} \quad \min_B \left\{ \min_{f \in \mathcal{H}_{\kappa_B}} \sum_{i=1}^d (Y^{(i)} - f(X^{(i)}))^2 + \lambda \|f\|^2_{\mathcal{H}_\mathcal{B}} \right\} \quad \text{”}$$

Hard to solve numerically...

Accounting for **anisotropy** in the kernel

Consider the modified Gaussian RBF

$$\kappa_B(x, x') = \exp\left(-\frac{1}{2\ell^2}\|B^\top(x - x')\|^2\right)$$

for some matrix $B \in \mathbb{R}^{d \times m}$ to determine...

$$\boxed{f \in \mathcal{H}_{\kappa_B} \qquad \Rightarrow \qquad f(x) = f(B^\top x)}$$

- **[Fukumizu 2004]**: optimize over $B \in \mathbb{R}^{d \times m}$

$$\text{“} \quad \min_B \left\{ \min_{f \in \mathcal{H}_{\kappa_B}} \sum_{i=1}^{d} (Y^{(i)} - f(X^{(i)}))^2 + \lambda \|f\|^2_{\mathcal{H}_\mathcal{B}} \right\} \quad \text{”}$$

  Hard to solve numerically...

- **[Samarov 1993]**: if $u(X) = \mathbb{E}[Y|X]$ is of the form $u(X) = f(B^\top x)$ then $\nabla u(x) = B \nabla f(B^\top x)$ so that

$$H = \mathbb{E}[\nabla u(X) \nabla u(X)^\top] = B \Big[\text{Some Matrix}\Big] B^\top$$

  Crazy simple: $\mathrm{eig}(H) \to B$, but requires the gradients u

Accounting for **anisotropy** in the kernel

Consider the modified Gaussian RBF

$$\kappa_B(x,x') = \exp\left(-\frac{1}{2\ell^2}\|B^\top(x-x')\|^2\right)$$

for some matrix $B \in \mathbb{R}^{d\times m}$ to determine...

$$\boxed{f \in \mathcal{H}_{\kappa_B} \qquad \Rightarrow \qquad f(x) = f(B^\top x)}$$

- **[Fukumizu 2004]**: optimize over $B \in \mathbb{R}^{d\times m}$

$$\text{“}\quad \min_B \left\{ \min_{f\in\mathcal{H}_{\kappa_B}} \sum_{i=1}^{d} (Y^{(i)} - f(X^{(i)}))^2 + \lambda \|f\|^2_{\mathcal{H}_{\mathcal{B}}} \right\} \quad\text{”}$$

  Hard to solve numerically...

- **[Samarov 1993]**: if $u(X) = \mathbb{E}[Y|X]$ is of the form $u(X) = f(B^\top x)$ then $\nabla u(x) = B\nabla f(B^\top x)$ so that

$$H = \mathbb{E}[\nabla u(X) \nabla u(X)^\top] = B\Big[\text{Some Matrix}\Big]B^\top$$

  Crazy simple: $\mathrm{eig}(H) \to B$, but requires the gradients u

- **[Fukumizu 2014]**: Gradient Kernel Dimension Reduction (gKDR)

$$\boxed{\text{replace } u(X) \text{ with } f^N(X) \in \mathcal{H}}$$

See how to optimize the kernel using gradients of $f^N(X)$

`gkdr_demo.ipynb`