Computational Information Games
A minitutorial Part I

Houman Owhadi

ICERM June 5, 2017


DARPA EQUiPS / AFOSR award no FA9550-16-1-0054
(Computational Information Games)
Probabilistic Numerical Methods

Statistical Inference approaches to numerical approximation and algorithm design


http://probabilistic-numerics.org/
http://oates.work/samsi
3 approaches to inference and to dealing with uncertainty

3 approaches to Numerical Approximation

- Worst Case
- Bayesian Average Probabilistic
- Adversarial Game


Deterministic zero sum game

Player I

Player II

How should I & II play the (repeated) game?
Worst case approach

II should play blue and lose 1 in the worst case

\[ \min_j \max_i G(i, j) \]
Worst case approach

\[
\begin{array}{c|c}
\text{Player II} & j \\
\hline
i & \begin{array}{c|c}
3 & -2 \\
-2 & 1 \\
\end{array}
\end{array}
\]

\[\max_i \min_j G(i, j)\]

I should play red and lose 2 in the worst case
No saddle point

Player I

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-2</td>
<td>1</td>
</tr>
</tbody>
</table>

Player II

\[
\max_i \min_j G(i, j) \neq \min_j \max_i G(i, j)
\]

Not an equilibrium for a repeated game
Average case (Bayesian) approach

Place a uniform prior on the choice of Player I

Player II

\[ p = \frac{1}{2} \]

Player I

\[ 1 - p = \frac{1}{2} \]

\[
\begin{array}{cc}
3 & -2 \\
-2 & 1 \\
\end{array}
\]

\[
\begin{array}{cc}
1/2 & -1/2 \\
\end{array}
\]

Average Loss of Player II

II Should always play blue

Not an equilibrium for a repeated game
Mixed strategy (repeated game) solution

Player II

\[
q = \frac{3}{8}, \quad 1 - q = \frac{5}{8}
\]

Player I

\[
\begin{array}{c c c c}
3 & 1 - p & -2 & 1 - p \\
\hline
p & -2 & 1 & 1 - p \\
1 - p & 1 & 1 & -2 \\
\end{array}
\]

Player II should play red with probability 3/8 and win 1/8 on average

Player I’s expected payoff =

\[
3pq + (1 - p)(1 - q) - 2p(1 - q) - 2q(1 - p)
\]

\[
= 1 - 3q + p(8q - 3) = -\frac{1}{8} \text{ for } q = \frac{3}{8}
\]
Mixed strategy (repeated game) solution

Player I

\[ p = \frac{3}{8} \]

\[ 1 - p = \frac{5}{8} \]

Player II

\[ q \]  \[ 1 - q \]

\begin{array}{cc}
3 & -2 \\
-2 & 1 \\
\end{array}

I should play red with probability 3/8 and lose 1/8 on average

Player I’s expected payoff

\[ = 3pq + (1 - p)(1 - q) - 2p(1 - q) - 2q(1 - p) \]

\[ = 1 - 3p + q(8p - 3) = -\frac{1}{8} \text{ for } q = \frac{3}{8} \]
Optimal strategies are mixed strategies

Optimal way to play is at random

Game theory

Optimal way to play is at random

John Von Neumann

\[ \min_q \max_p q_j p_i G(i, j) = \max_p \min_q q_j p_i G(i, j) \]
The optimal mixed strategy is determined by the loss matrix:

\[
q = \frac{3}{10}, \quad 1 - q = \frac{7}{10}
\]

Player II should play red with probability 3/10 and win 1/8 on average.

Player I’s expected payoff is:

\[
5pq + (1 - p)(1 - q) - 2p(1 - q) - 2q(1 - p)
\]

\[
= 1 - 3q + p(10q - 3) = -\frac{1}{8} \text{ for } q = \frac{3}{10}
\]
Pioneering work

[ Henri Poincaré. Calcul des probabilités. 1896. ]

[ A. V. Sul’din, Wiener measure and its applications to approximation methods. Matematika 1959 ]

[ A. Sard. Linear approximation. 1963. ]

[ G. S. Kimeldorf and G. Wahba. A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. 1970 ]

[ F.M. Larkin. Gaussian measure in Hilbert space and applications in numerical analysis. Rocky Mountain J. Math, 1972 ]

“ These concepts and techniques have attracted little attention among numerical analysts” (Larkin, 1972)
Bayesian Numerical Analysis

[ P. Diaconis. Bayesian numerical analysis. In Statistical decision theory and related topics, 1988 ]
[ Skilling, J. Bayesian solution of ordinary differential equations. 1992. ]
Information based complexity


[ Erich Novak and Henryk Woźniakowski, Tractability of Multivariate Problems, 2008-2010 ]
Compute

\[ \int_0^1 f(x) \, dx \]

**Numerical Analysis Approach**

Find a good quadrature rule for the numerical integration of \( f \)

[ P. Diaconis. Bayesian numerical analysis. In Statistical decision theory and related topics, 1988 ]
\[ f(x) = \exp \left( \cosh \left( \frac{x^2 + \sin(x)}{3 + \cos(x^3)} \right) \right) \]

**Compute**

\[
\int_0^1 f(x) \, dx
\]

**Bayesian Approach**

- Put a prior in \( \mathcal{C}([0, 1]) \)
- Calculate \( f \) at \( x_1, \ldots, x_n \)
- Compute

\[
\mathbb{E} \left[ \int_0^1 f(x) \, dx \bigg| f(x_1), \ldots, f(x_n) \right]
\]
E.g.

Assume $f(t) = \xi + B_t$

$\mathcal{N}(0, 1) \quad \text{B.M.}$

$$\mathbb{E}\left[ f(x) \Big| f(x_1), \ldots, f(x_n) \right] = \sum_i f(x_i)\psi_i(x)$$

Piecewise linear interpolation of $f$

$$\mathbb{E}\left[ \int_0^1 f(x) \, dx \Big| f(x_1), \ldots, f(x_n) \right] \rightarrow \text{Trapezoidal quadrature rule}$$
E.g.

Assume \( f(t) = \xi + \int_0^t B_s \, ds \)

\[ \mathcal{N}(0, 1) \quad \text{B.M.} \]

\[ \mathbb{E}\left[ f(x) \, \bigg| \, f(x_1), \ldots, f(x_n) \right] \rightarrow \text{Cubic spline interpolant} \]

**E.g.**

Integrate B.M. \( k \) times

Splines of order \( 2k + 1 \)

Similar link between PDEs and Bayesian Inference?

\[
\begin{aligned}
\begin{cases}
- \text{div}(a \nabla u) = g, & x \in \Omega, \\
\quad u = 0, & x \in \partial\Omega,
\end{cases}
\end{aligned}
\]

\(\Omega \subset \mathbb{R}^d\) \hspace{1cm} \partial\Omega \text{ is piec. Lip.} \\
a \text{ unif. ell.} \quad a_{i,j} \in L^\infty(\Omega) \\
g \in L^2(\Omega)

Approximate the solution space of (1) with a finite dimensional space
Numerical Homogenization Approach

Work hard to find good basis functions

**Harmonic Coordinates**  Babuska, Caloz, Osborn, 1994
                        Kozlov, 1979  Allaire Brizzi 2005; Owhadi, Zhang 2005

**MsFEM**            [Hou, Wu: 1997]; [Efendiev, Hou, Wu: 1999]
                     [Fish - Wagiman, 1993]  [Chung-Efendiev-Hou, JCP 2016]

**Variational Multiscale Method, Orthogonal decomposition**
                     [Hughes, Feijóo, Mazzei, Quincy. 1998]
                     [Malqvist-Peterseim 2012]  Local Orthogonal Decomposition

**Projection based method**  Nolen, Papanicolaou, Pironneau, 2008

**HMM**           Engquist, E, Abdulle, Runborg, Schwab, et Al. 2003-...

**Flux norm**    Berlyand, Owhadi 2010; Symes 2012

**Harmonic continuation**  [Babuska-Lipton 2010]
Bayesian Approach

\[
\begin{cases}
- \operatorname{div}(a \nabla u) = g, & x \in \Omega, \\
u = 0, & x \in \partial \Omega,
\end{cases}
\]

Proposition

- Put a prior on \( g \)
- Compute \( \mathbb{E}[u(x) | \text{finite no of observations}] \)
Bayesian approach

Replace $g$ by $\xi$

\[
\begin{cases}
- \text{div}(a \nabla v) = \xi, & \Omega, \\
u = 0, & \partial \Omega.
\end{cases}
\]

$\xi$: White noise

Gaussian field with covariance function $\Lambda(x, y) = \delta(x - y)$

$\Leftrightarrow \forall f \in L^2(\Omega), \int_{\Omega} f(x)\xi(x) \, dx$ is $\mathcal{N}(0, \|f\|_{L^2(\Omega)}^2)$


Let \( x_1, \ldots, x_N \in \Omega \) 

**Theorem**

\[
\mathbb{E}\left[ v(x) \bigg| v(x_1), \ldots, v(x_N) \right] = \sum_{i=1}^{N} v(x_i) \psi_i(x)
\]

\( a = I_d \quad \psi_i: \text{Polyharmonic splines} \)

[Harder-Desmarais, 1972]


\( a_{i,j} \in L^\infty(\Omega) \quad \psi_i: \text{Rough Polyharmonic splines} \)

[Owhadi-Zhang-Berlyand 2013]
Standard deviation of the statistical error bounds/controls the worst case error

\[ (v(x)|v(x_1), \ldots, v(x_N)) \sim \mathcal{N} \left( \sum_{i=1}^{N} v(x_i)\psi_i(x), \sigma^2(x) \right) \]

\(\sigma^2(x)\): Kriging function (geostatistics)


\(\sigma^2(x)\): Power function (radial basis function interpolation)


**Theorem**

\[ |u(x) - \sum_{i=1}^{N} u(x_i)\psi_i(x)| \leq \sigma(x) \| g \|_{L^2(\Omega)} \]
The Bayesian approach leads to old and new quadrature rules.

Statistical errors seem to imply/control deterministic worst case errors

Questions

• Why does it work?
• How far can we push it?
• What are its limitations?
• How can we make sense of the process of randomizing a known function?
$B_1 \xrightarrow{\mathcal{L}} B_2$

$u \xrightarrow{} g$

**Direct Problem**

Given $u$ find $g$

**Inverse Problem**

Given $g$ find $u$

$u$ and $g$ live in infinite dimensional spaces

Direct computation is not possible

Inverse Problem

Reduced operator

\[ \mathcal{L} \]

\[ \mathcal{B}_1 \xrightarrow{\mathcal{L}} \mathcal{B}_2 \]

Numerical implementation requires computation with partial information.

\[ \phi_1, \ldots, \phi_m \in \mathcal{B}_1^* \]

\[ u_m = ([\phi_1, u], \ldots, [\phi_m, u]) \]

\[ u_m \in \mathbb{R}^m \text{ Missing information} \]

\[ u \in \mathcal{B}_1 \]
**Fast Solvers**

**Multigrid Methods**


**Multiresolution/Wavelet based methods**


**Robust/Algebraic multigrid**


**Stabilized Hierarchical bases, Multilevel preconditioners**

[Vassilevski - Wang, 1997, 1998] [Chow - Vassilevski, 2003]
[Panayot - Vassilevski, 1997] [Aksoylu- Holst, 2010]

**Low rank matrix decomposition methods**

Fast Multipole Method: [Greengard and Rokhlin, 1987]
Hierarchical Matrix Method: [Hackbusch et al., 2002] [Bebendorf, 2008]:
Common theme between these methods

Computation is done with partial information over hierarchies of levels of complexity

\[ u \in \mathcal{B} \]

\[ u_m = ([\phi_1, u], \ldots, [\phi_m, u]) \]

\[ \phi_1, \ldots, \phi_m \in \mathcal{B}^* \]

To compute fast we need to compute with partial information
The process of discovery of interpolation operators is based on intuition, brilliant insight, and guesswork.

\[ u_m \xrightarrow{\text{Missing information}} u \]

**Problem**

Given \([\phi_1, u], \ldots, [\phi_m, u]\) recover \(u\)

This is one entry point for statistical inference into Numerical analysis and algorithm design.
A simple approximation problem

Approximate

\[ x \in \mathbb{R}^n \]

Based on the information that

\[ \Phi x = y \]

\( \Phi \): Known \( m \times n \) rank \( m \) matrix \( (m < n) \)

\( y \): Known element of \( \mathbb{R}^m \)

\( \nu(y) \)  Your approximation
Worst case approach (Optimal Recovery)

**Problem** \[ \| \cdot \| : \text{Quadratic norm on } \mathbb{R}^n \]

Find \( \nu : \mathbb{R}^m \to \mathbb{R}^n \) minimizing worst case error

\[
\inf_{\nu} \sup_{x \in \mathbb{R}^n} \frac{\| x - \nu(\Phi x) \|}{\| x \|}
\]

---


$v(y)$ is the minimizer of

\[
\begin{cases}
\text{Minimize } \|w\| \\
\text{Subject to } w \in \mathbb{R}^n \text{ and } \Phi w = y
\end{cases}
\]

$v(y) = \sum_i y_i \psi_i \quad \psi_i: \text{Optimal recovery splines}$


**Average case approach (IBC)**

\[ \| \cdot \|: \text{Quadratic norm on } \mathbb{R}^n \]

\[ \mu: \text{Measure of probability on } \mathbb{R}^n \text{ s.t. } \int \|x\|^2 \mu(dx) < \infty \]

\[ \mathcal{E}(v) = \int \|x - v(\Phi x)\|^2 \mu(dx): \text{Average error} \]

**Problem**

Find \( v: \mathbb{R}^m \rightarrow \mathbb{R}^n \) minimizing average error

---


Solution

Has a natural Bayesian interpretation

\[ \mu = \mathcal{N}(0, C) \quad \leftrightarrow \quad v(y) = \mathbb{E}_{x \sim \mu} [x | \Phi x = y] \]

\( v(y) \) is the minimizer of

\[
\begin{aligned}
\text{Minimize} & \quad w^T C^{-1} w \\
\text{Subject to} & \quad w \in \mathbb{R}^n \text{ and } \Phi w = y
\end{aligned}
\]

If \( \|x\|^2 = x^T C^{-1} x \) and \( \mu = \mathcal{N}(0, C) \) then average case solution = worst case solution

Adversarial game approach

**Player I**

Chooses $x \in \mathbb{R}^n$

- **Max**

**Player II**

Sees $y = \Phi x$

Chooses $v(y)$

- **Min**

\[
\frac{\| x - v(\Phi x) \|^2}{\| x \|^2}
\]

Loss function

\[ \mathcal{E}(x, v) = \frac{\|x - v(\Phi x)\|^2}{\|x\|^2} \]

Player I

\[ \max_x \min_v \mathcal{E}(x, v) = 0 \]

Player II

\[ \min_v \max_x \mathcal{E}(x, v) \neq 0 \]

No saddle point of pure strategies
Randomized strategy for player I

**Player I**

Chooses $\mu \in \mathcal{P}(\mathbb{R}^n)$

Samples $x \sim \mu$

Max

**Player II**

Sees $y = \Phi x$

Chooses $v(y)$

$$\max \quad \frac{\int_{\mathbb{R}^n} \| x - v(\Phi x) \|^2 \mu(dx)}{\int_{\mathbb{R}^n} \| x \|^2 \mu(dx)}$$

Loss function

\[ \mathcal{E}(\mu, \nu) = \frac{\int_{\mathbb{R}^n} \| x - \nu(\Phi x) \|^2 \mu(dx)}{\int_{\mathbb{R}^n} \| x \|^2 \mu(dx)} \]

Saddle point

\[ \max_\mu \min_\nu \mathcal{E}(\mu, \nu) = \min_\nu \max_\mu \mathcal{E}(\mu, \nu) \]

\[ \exists \mu^\dagger, \nu^\dagger \]

\[ \mathcal{E}(\mu, \nu^\dagger) \leq \mathcal{E}(\mu^\dagger, \nu^\dagger) \text{ for all } \mu \]

\[ \mathcal{E}(\mu^\dagger, \nu) \geq \mathcal{E}(\mu^\dagger, \nu^\dagger) \text{ for all } \nu \]
Canonical Gaussian field

\[ \|x\|^2 := x^T A x \]

\(A: n \times n\) symmetric positive definite matrix

\(\xi: \text{Canonical Gaussian field on } (\mathbb{R}^n, \| \cdot \|)\)

Density function

\[ f(x) = \frac{e^{-\frac{\|x\|^2}{2}}}{C} \]

\(\xi \sim \mathcal{N}(0, A^{-1})\)

For \(z \in \mathbb{R}^n\), \(z^T \xi \sim \mathcal{N}(0, \|z\|^2_*)\)

\[ \|z\|_* = \sup_{x \in \mathbb{R}^n} \frac{z^T x}{\|x\|} = z^T A^{-1} z \]
Equilibrium saddle point

Player I
\[ \mu^\dagger \leftrightarrow x \sim \xi - \mathbb{E}[\xi | \Phi \xi] \]

Player II
\[ \nu^\dagger (y) = \mathbb{E}[\xi | \Phi \xi = y] \]

\[ \mathbb{R}^m \xrightarrow{y} \mathbb{R}^n \xrightarrow{x} \]
The optimal bet of Player II is Bayesian

Complete Class Theorem

Statistical decision theory


The game theoretic solution is equal to the worst case solution

\[ v^\dagger(y) = \mathbb{E}[\xi \mid \Phi \xi = y] \]

\( \xi \) has density \( e^{-\frac{\|w\|^2}{2C}} \)

\[ \|w\| = \text{constant} \]

\[ \Phi w = y \]

\( v^\dagger(y) \) is the minimizer of

\[
\begin{align*}
\text{Minimize} & \quad \|w\| \\
\text{Subject to} & \quad w \in \mathbb{R}^n \text{ and } \Phi w = y
\end{align*}
\]
$(\mathcal{B}, \| \cdot \|)$: separable Banach space

$\| \cdot \|$: Quadratic norm

$\| u \|^2 := [\mathcal{T} u, u]$

$\mathcal{T}$: Symmetric positive continuous linear bijection

$\mathcal{B} \xrightarrow{\mathcal{T}} \mathcal{B}^*$

For $u, v \in \mathcal{B}$,

- $[\mathcal{T} u, v] = [\mathcal{T} v, u]$
- $[\mathcal{T} u, u] \geq 0$
Examples

\[ \mathcal{B} := \mathbb{R}^N \quad \|x\|^2 := x^T A x \]

A: \( N \times N \) symmetric positive definite matrix

\[ \mathcal{B} := H^s_0(\Omega) \quad \|u\|^2 := \int_\Omega u \mathcal{L} u \]

\( \mathcal{L} \): arbitrary symmetric, positive, continuous linear bijection

\( (H^s_0(\Omega), \| \cdot \|_{H^s_0(\Omega)}) \xrightarrow{\mathcal{L}} (H^{-s}(\Omega), \| \cdot \|_{H^{-s}(\Omega)}) \)
Canonical Gaussian field

\[ \| x \|^2 := x^T A x \]

\( A: n \times n \) symmetric positive definite matrix

\( \xi: \) Canonical Gaussian field on \( (\mathbb{R}^n, \| \cdot \|) \)

\[ f(x) = \frac{e^{-\frac{\| x \|^2}{2}}}{C} \]

\( \xi \sim \mathcal{N}(0, A^{-1}) \)

For \( z \in \mathbb{R}^n \), \( z^T \xi \sim \mathcal{N}(0, \| z \|^2) \)

\[ \| z \|_* = \sup_{x \in \mathbb{R}^n} \frac{z^T x}{\| x \|} = z^T A^{-1} z \]
(\mathcal{B}, \| \cdot \|): \text{ separable Banach space}

\| \cdot \|: \text{ Quadratic norm}

\xi: \mathcal{B}^* \rightarrow \mathcal{H}

\phi \rightarrow [\phi, \xi] \sim \mathcal{N}(0, \| \phi \|_2^2)

\xi: \text{ Linear isometry mapping } \mathcal{B}^* \text{ to a Gaussian Space}

\mathbb{E} \left[ [\varphi, \xi] [\phi, \xi] \right] = \langle \varphi, \phi \rangle_* \quad \| \phi \|_* := \sup_{\nu \in \mathcal{B}} \frac{[\phi, \nu]}{\| \nu \|}
Canonical Gaussian field

\[ \|u\|^2 := [\mathcal{T} u, u] \]

\( \mathcal{T} \): Symmetric positive continuous linear bijection

\[ \mathcal{B} \xrightarrow{\mathcal{T}} \mathcal{B}^* \]

\( \xi \sim \mathcal{N}(0, \mathcal{T}^{-1}) \)

For \( \varphi, \phi \in \mathcal{B}^* \)

\[ \mathbb{E}[\langle \varphi, \xi \rangle \langle \phi, \xi \rangle] = [\varphi, \mathcal{T}^{-1} \phi] \]
\[ \mathcal{B} := \mathbb{R}^N \quad \text{and} \quad \|x\|^2 := x^T A x \]

A: \( N \times N \) symmetric positive definite matrix

\[
\xi = \mathcal{N}(0, A^{-1})
\]

\[ \mathcal{B} := H_0^s(\Omega) \quad \|u\|^2 := [\mathcal{L}u, u] \]

\( \mathcal{L} \): arbitrary symmetric, positive, continuous linear bijection

\[
(H_0^s(\Omega), \| \cdot \|_{H_0^s(\Omega)}) \xrightarrow{\mathcal{L}} (H^{-s}(\Omega), \| \cdot \|_{H^{-s}(\Omega)})
\]

\[
\int_{\Omega} \phi(x) \xi(x) \, dx \sim \mathcal{N}(0, \int_{\Omega^2} \phi(x) G(x, y) \phi(y) \, dx \, dy)
\]
The recovery problem at the core of Algorithm Design and Numerical Analysis

To compute fast we need to compute with partial information

\[ u \in \mathcal{B} \]

**Restriction** \[ \downarrow \]
\[ u_m = ([\phi_1, u], \ldots, [\phi_m, u]) \]
\[ \phi_1, \ldots, \phi_m \in \mathcal{B}^* \]

**Interpolation** \[ \uparrow \]

\[ u_m \xrightarrow{\text{Missing information}} u \]

**Problem**

Given \(([\phi_1, u], \ldots, [\phi_m, u])\) recover \(u\)
\[ \phi_1, \ldots, \phi_m \in \mathcal{B}^* \]

**Player I**

Chooses \( u \in \mathcal{B} \)

\[ \text{Max} \]

\[ \frac{\| u - v([\phi_1, u], \ldots, [\phi_m, u]) \|^2}{\| u \|^2} \]

**Player II**

Sees \( y = ([\phi_1, u], \ldots, [\phi_m, u]) \)

Chooses \( v(y) \in \mathcal{B} \)

\[ \text{Min} \]
Examples

Player I

Chooses $x \in \mathbb{R}^N$

Sees $(\phi_1^T x, \ldots, \phi_m^T x)$

Chooses $v(\phi_1^T x, \ldots, \phi_m^T x) \in \mathbb{R}^N$

Player II

$\mathcal{B} := \mathbb{R}^N$

Player I

Chooses $u \in H_0^s(\Omega)$

Sees $(\int_{\tau_1} u, \ldots, \int_{\tau_m} u)$

Chooses $v(\int_{\tau_1} u, \ldots, \int_{\tau_m} u) \in H_0^s(\Omega)$

Player II

$\mathcal{B} := H_0^s(\Omega)$
Loss function

\[ \mathcal{E}(u, v) = \left\| \frac{u - v([\phi_1, u], \ldots, [\phi_m, u])}{\|u\|^2} \right\|^2 \]

Player I

\[ \max_u \min_v \mathcal{E}(u, v) = 0 \]

Player II

\[ \min_v \max_u \mathcal{E}(u, v) \neq 0 \]

No saddle point of pure strategies
Randomized strategy for player I

\[ \phi_1, \ldots, \phi_m \in \mathcal{B}^* \]

**Player I**

Chooses \( \mu \in \mathcal{P}(\mathcal{B}) \)

Samples \( u \sim \mu \)

**Player II**

Sees \( y = ([\phi_1, u], \ldots, [\phi_m, u]) \)

Chooses \( v(y) \in \mathcal{B} \)

\[
\begin{align*}
\max & \int \left\| u - v([\phi_1, u], \ldots, [\phi_m, u]) \right\|^2 \mu(du) \\
\min & \int \|u\|^2 \mu(du)
\end{align*}
\]
Loss function

\[ \mathcal{E}(\mu, v) = \frac{\int \left\| u - v([\phi_1, u], ..., [\phi_m, u]) \right\|^2 \mu(du)}{\int \|u\|^2 \mu(du)} \]

Theorem

\[ \sup_{\mu} \inf_{v} \mathcal{E}(\mu, v) = \inf_{v} \sup_{\mu} \mathcal{E}(\mu, v) \]

But

But no saddle point if \( \dim(\mathcal{B}) = \infty \)

No \( \mu^* \in \mathcal{P}(\mathcal{B}) \) is achieving the \( \sup_{\mu} \)}
Loss function

\[ \mathcal{E}(\mu, \nu) = \frac{\int \| u - \nu([\phi_1, u], \ldots, [\phi_m, u]) \|^2 \mu(du)}{\int \| u \|^2 \mu(du)} \]

Theorem

\[ \max_\mu \min_\nu \mathcal{E}(\mu, \nu) = \min_\nu \max_\mu \mathcal{E}(\mu, \nu) \]

Definition \( \mu^*, \nu^* \) is a \( \epsilon \) saddle point iff

\[ \mathcal{E}(\mu, \nu^*) \leq \mathcal{E}(\mu^*, \nu^*) + \epsilon \text{ for all } \mu \]

\[ \mathcal{E}(\mu^*, \nu) \geq \mathcal{E}(\mu^*, \nu^*) - \epsilon \text{ for all } \nu \]
Theorem

\[ \exists \nu^\dagger : \mathbb{R}^m \rightarrow \mathcal{B} \]
\[ \exists \mu_n \in \mathcal{P}(\mathcal{B}) \text{ indexed by } n \in \mathbb{N}^* \]
\[ \exists \mu^\dagger \text{ cylinder measure on } \mathcal{B} \]

s.t. for all \( \mu, \nu \)

\[ \mathcal{E}(\mu, \nu^\dagger) - \frac{1}{n} \leq \mathcal{E}(\mu_n, \nu^\dagger) \leq \mathcal{E}(\mu_n, \nu) + \frac{1}{n} \]

and

\[ \mu_n \xrightarrow{w} \mu^\dagger \quad \text{as} \quad n \to \infty \]

For all \( s \)

\[ \mathbb{E}_{u \sim \mu_n} \left[ \left( [\varphi_1, u], \ldots, [\varphi_s, u] \right) \right] \xrightarrow{n \to \infty} \mathbb{E}_{u \sim \mu^\dagger} \left[ \left( [\varphi_1, u], \ldots, [\varphi_s, u] \right) \right] \]
Theorem The optimal mixed strategy for Player I is

$$\mu^\dagger \iff u \sim \xi - \mathbb{E}[\xi \mid ([\phi_i, \xi])_{i \in \mathcal{I}}]$$

The optimal strategy for Player II is

$$\nu([\phi_1, u], \ldots, [\phi_m, u]) = \mathbb{E}[\xi \mid [\phi_i, \xi] = [\phi_i, u] \text{ for } i \in \mathcal{I}]$$
\[ v(y) = \mathbb{E}[\xi \mid [\phi_i, \xi] = y_i \text{ for } i \in \mathcal{I}] \]

\[ \|w\| = \text{constant} \]

\[ \Phi w = y \]

\[ v(y): \text{ Minimizer of} \]

\[
\begin{cases}
\text{Minimize } \|w\| \\
\text{Subject to } w \in \mathcal{B} \text{ and } [\phi_i, w] = y_i \text{ for } i \in \mathcal{I}
\end{cases}
\]
Game theoretic solution = Worst case solution

\[ v(y) = \mathbb{E} \left[ \xi \mid [\phi_i, \xi] = y_i \text{ for } i \in I \right] \]

Optimal Recovery Solution

\[ v(y) : \text{Minimizer of} \]

\[ \inf_v \sup_{u \in B} \left| \left| u - v([\phi_1, u], \ldots, [\phi_m, u]) \right| \right|^2 / \| u \|^2 \]


Optimal bet of player II

\[ u^* = \mathbb{E} [\xi \mid [\phi_i, \xi] = [\phi_i, u] \text{ for } i \in I] \]

Gamblets

\[ u^* = \sum_i [\phi_i, u] \psi_i \]

\[ \psi_i \in B \]

\[ \psi_i = \mathbb{E} [\xi \mid [\phi_j, \xi] = \delta_{i,j} \text{ for } j \in I] \]
Gamblets = Optimal Recovery Splines

\[ \psi_i = \mathbb{E}[\xi \mid [\phi_j, \xi] = \delta_{i,j} \text{ for } j \in \mathcal{I}] \]

Optimal Recovery Splines

\[ \psi_i \text{ is the minimizer of} \]

\[
\left\{
\begin{array}{l}
\text{Minimize } \|w\| \\
\text{Subject to } w \in \mathcal{B} \text{ and } [\phi_j, w] = \delta_{i,j} \text{ for } j \in \mathcal{I}
\end{array}
\right.
\]


\( \psi_i \) is the minimizer of
\[
\begin{aligned}
\text{Minimize } & \|w\| \\
\text{Subject to } & w \in \mathcal{B} \text{ and } [\phi_j, w] = \delta_{i,j} \text{ for } j \in \mathcal{I}
\end{aligned}
\]

\( \phi_i \) is the minimizer of
\[
\begin{aligned}
\text{Minimize } & ||\phi||_* \\
\text{Subject to } & \phi \in \mathcal{B}^* \text{ and } [\phi, \psi_j] = \delta_{i,j} \text{ for } j \in \mathcal{I}
\end{aligned}
\]
Example

\[ \mathcal{B} := H^1_0(\Omega) \quad \|u\|^2 := \int_\Omega (\nabla u)^T a \nabla u \]

\[ \mathcal{T} = - \text{div}(a \nabla \cdot) \]

\[ \xi \sim \mathcal{N}(0, \mathcal{T}^{-1}) \]

\[ \begin{cases} 
- \text{div}(a \nabla u) = g, & x \in \Omega, \\
\quad u = 0, & x \in \partial \Omega, 
\end{cases} \]
Your best bet on the value of $u$ given the information that
\( \int_{\tau_i} u = 1 \) and \( \int_{\tau_j} u = 0 \) for \( j \neq i \).
Example

\[ \mathcal{B} := H_0^1(0, 1) \]

\[ \|u\|^2 = \int_0^1 \left( \frac{du}{dx} \right)^2 \, dx \]

\[ \tau = -\frac{d^2}{(dx)^2} \]

\[ \xi: \text{Brownian bridge} \]

\[ \psi_i(x) \]

\[ \phi_i(x) = \delta(x - x_i) \]
Example

\[ \mathcal{B} := H^1(0, 1) \]

\[ \|u\|^2 = a(u(0))^2 + b \int_0^1 \left( \frac{du}{dx} \right)^2 \, dx \]

\[ \xi_t = \alpha \mathcal{N}(0, 1) + \beta \mathcal{B}_t \]

\[ \psi_i(x) \quad \phi_i(x) = \delta(x - x_i) \]

\[ \mathbb{E}[\xi(x) \mid \xi(x_1) = f(x_1), \ldots, \xi(x_n) = f(x_n)] \quad \text{Piecewise linear interpolation of } f \]
\[ \mathcal{B} := H^1_0(\Omega) \cap H^2(\Omega) \]

\[ \mathcal{T} = \Delta^2 \quad \xi \sim \mathcal{N}(0, \mathcal{T}^{-1}) \]

\[ \phi_i(x) = \delta(x - x_i) \]

\[ \psi_i: \text{Polyharmonic splines} \]

[Harder-Desmarais, 1972] [Duchon 1976, 1977, 1978]
Example

\[ \mathcal{B} := \{ u \in H^1_0(\Omega) \mid \int_\Omega |\text{div}(a \nabla u)|^2 < \infty \} \]

\[ |u|^2 := \int_\Omega |\text{div}(a \nabla u)|^2 \quad a_{i,j} \in L^\infty(\Omega) \]

\[ \mathcal{T} = (-\text{div}(a \nabla \cdot))^2 \]

\[ \phi_i(x) = \delta(x - x_i) \]

\[ \psi_i: \text{Rough Polyharmonic splines} \]

[Owhadi-Zhang-Berlyand 2013]
Example

\[ \mathcal{B} := H^1_0(\Omega) \]

\[ \mathcal{T} = - \text{div}(a \nabla \cdot) \]

\[ \phi_i = \sum_j M^{-1}_{i,j} \varphi_j \]

\[ \|u\|^2 := \int_\Omega (\nabla u)^T a (\nabla u)^T \]

\[ \psi_i: \text{ LOD basis} \]

Example

\[ \mathcal{B} := H^1_0(\Omega) \]
\[ \mathcal{T} = -\text{div}(a \nabla \cdot) \]

\[ \phi_i = 1_{\mathcal{T}_i} \]

\[ \|u\|^2 := \int_{\Omega} (\nabla u)^T a (\nabla u)^T \]

Summary

• Bayesian numerical analysis "works" because Gaussian priors form the optimal class of priors when losses are defined using quadratic norms and measurements are linear.
• The game theoretic solution is equal to the classical worst case optimal recovery solution under above questions.
• The canonical Gaussian field contains all the required information to bridge scales/levels of complexity in numerical approximation and it does not depend on the linear measurements.

Questions

• Does the canonical Gaussian field remain optimal (or near optimal) beyond average relative errors (e.g. rare events/large deviations) or when measurements are not linear. This is a fundamental question if probabilistic numerical errors are to be merged with model errors in a unified Bayesian framework.
• What are the properties of gamblets?
• Can the game theoretic approach help us solve known open problems in numerical analysis and algorithm design?
Thank you


- Multigrid with gambles. L. Zhang and H. Owhadi, 2017


DARPA EQUiPS / AFOSR award no FA9550-16-1-0054 (Computational Information Games)