Lecture 10

Nonuniqueness
and
Localized Averages

Syllabus

Lecture 01	Describing Inverse Problems
Lecture 02	Probability and Measurement Error, Part 1
Lecture 03	Probability and Measurement Error, Part 2
Lecture 04	The L_2 Norm and Simple Least Squares
Lecture 05	A Priori Information and Weighted Least Squared
Lecture 06	Resolution and Generalized Inverses
Lecture 07	Backus-Gilbert Inverse and the Trade Off of Resolution and Variance
Lecture 08	The Principle of Maximum Likelihood
Lecture 09	Inexact Theories
Lecture 10	**Nonuniqueness and Localized Averages**
Lecture 11	Vector Spaces and Singular Value Decomposition
Lecture 12	Equality and Inequality Constraints
Lecture 13	L_1 , L_∞ Norm Problems and Linear Programming
Lecture 14	Nonlinear Problems: Grid and Monte Carlo Searches
Lecture 15	Nonlinear Problems: Newton's Method
Lecture 16	Nonlinear Problems: Simulated Annealing and Bootstrap Confidence Intervals
Lecture 17	Factor Analysis
Lecture 18	Varimax Factors, Empirical Orthogonal Functions
Lecture 19	Backus-Gilbert Theory for Continuous Problems; Radon's Problem
Lecture 20	Linear Operators and Their Adjoints
Lecture 21	Fréchet Derivatives
Lecture 22	Exemplary Inverse Problems, incl. Filter Design
Lecture 23	Exemplary Inverse Problems, incl. Earthquake Location
Lecture 24	Exemplary Inverse Problems, incl. Vibrational Problems

Purpose of the Lecture

Show that null vectors are the source of nonuniqueness

Show why some localized averages of model parameters are unique while others aren't

Show how nonunique averages can be bounded using prior information on the bounds of the underlying model parameters

Introduce the Linear Programming Problem

Part 1

null vectors as the source of nonuniqueness in linear inverse problems

suppose two different solutions exactly satisfy the same data

$$\mathbf{G}\mathbf{m}^{(1)} = \mathbf{d}$$
$$\mathbf{G}\mathbf{m}^{(2)} = \mathbf{d}$$

since there are two the solution is nonunique then the difference between the solutions satisfies

$$\mathbf{G}\left(\mathbf{m}^{(1)} - \mathbf{m}^{(2)}\right) = \mathbf{0}$$

the quantity

$$\mathbf{m}^{\text{null}} = \mathbf{m}^{(1)} - \mathbf{m}^{(2)}$$

is called a *null vector*

it satisfies

$$\mathbf{G}\,\mathbf{m}^{\text{null}} = \mathbf{0}$$

an inverse problem can have more than one null vector

$$\mathbf{m}^{\text{null}(1)} \quad \mathbf{m}^{\text{null}(2)} \quad \mathbf{m}^{\text{null}(3)} \ldots$$

any linear combination of null vectors is a null vector

$$\alpha\mathbf{m}^{\text{null}(1)} + \beta\mathbf{m}^{\text{null}(2)} + \gamma\mathbf{m}^{\text{null}(3)}$$

is a null vector for any α, β, γ

suppose that a particular choice of model parameters

$$\mathbf{m}^{\text{par}}$$

satisfies

$$\mathbf{G}\,\mathbf{m}^{\text{par}} = \mathbf{d}^{\text{obs}}$$

with error E

then

$$\mathbf{m}^{\text{gen}} = \mathbf{m}^{\text{par}} + \sum_{i=1}^{q} \alpha_i \, \mathbf{m}^{\text{null (i)}}$$

has the same error E
for any choice of α_i

since

$$\mathbf{e} = \mathbf{d}^{obs} - \mathbf{Gm}^{gen} = \mathbf{d}^{obs} - \mathbf{Gm}^{par} + \Sigma_i\ \alpha_i\ 0$$

since

$$\mathbf{m}^{\text{gen}} = \mathbf{m}^{\text{par}} + \sum_{i=1}^{q} \alpha_i \, \mathbf{m}^{\text{null (i)}}$$

since α_i is arbitrary
the solution is nonunique

hence

an inverse problem is nonunique

if it has null vectors

example
consider the inverse problem

$$\mathbf{Gm} = \begin{bmatrix} \tfrac{1}{4} & \tfrac{1}{4} & \tfrac{1}{4} & \tfrac{1}{4} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix} = [d_1]$$

a solution with zero error is
$\mathbf{m}^{\mathrm{par}}=[d_1, d_1, d_1, d_1]^T$

the null vectors are easy to work out

$$\mathbf{m}^{\text{null}\,(1)} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{m}^{\text{null}\,(2)} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{m}^{\text{null}\,(3)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

note that $\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ times any of these vectors is zero

the general solution to the inverse problem is

$$\mathbf{m}^{\text{gen}} = \begin{bmatrix} d_1 \\ d_1 \\ d_1 \\ d_1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

Part 2

Why some localized averages are unique while others aren't

let's denote a weighted average of the model parameters as

$$\langle m \rangle = \mathbf{a}^T \mathbf{m}$$

where **a** is the vector of weights

let's denote a weighted average of the model parameters as

$$<m> = \mathbf{a}^T \mathbf{m}$$

where **a** is the vector of weights

a may or may not be "localized"

examples

$$\mathbf{a} = [0.25,\ 0.25,\ 0.25,\ 0.25]^{\mathrm{T}}$$

not localized

$$\mathbf{a} = [0.\ 90,\ 0.07,\ 0.02,\ 0.01]^{\mathrm{T}}$$

localized near m_1

now compute the average of the general solution

$$\langle \mathrm{m} \rangle = \mathbf{a}^{\mathrm{T}} \mathbf{m}^{\mathrm{gen}} = \mathbf{a}^{\mathrm{T}} \mathbf{m}^{\mathrm{par}} + \sum_{i=1}^{q} \alpha_i \, \mathbf{a}^{\mathrm{T}} \mathbf{m}^{\mathrm{null}\,(\mathrm{i})}$$

now compute the average of the general solution

$$\langle \mathrm{m} \rangle = \mathbf{a}^{\mathrm{T}}\mathbf{m}^{\mathrm{gen}} = \mathbf{a}^{\mathrm{T}}\mathbf{m}^{\mathrm{par}} + \sum_{i=1}^{q} \alpha_i \, \mathbf{a}^{\mathrm{T}}\mathbf{m}^{\mathrm{null}\ (\mathrm{i})}$$

if this term is zero for all *i*,
then $<m>$ does not depend on α_i,
so average is unique

an average $\langle m \rangle = \mathbf{a}^T \mathbf{m}$ is unique

if the average of all the null vectors

is zero

if we just pick an average
out of the hat
because we like it ... its nicely localized

chances are that it will not zero all the null vectors

so the average will not be unique

relationship to model resolution **R**

$$\mathbf{m}^{\text{est}} = \mathbf{R}\mathbf{m}^{\text{true}} = \mathbf{G}^{-g}\mathbf{G}\mathbf{m}^{\text{true}} \quad \text{or} \quad m_i^{est} = \sum_{j=1}^{M} R_{ij} m_j^{true} = \sum_{j=1}^{M}\sum_{k=1}^{N} G_{ik}^{-g} G_{kj}\, m_j^{true}$$

$$m_i^{est} = \langle m \rangle^{(i)} = \sum_{j=1}^{M} a_j^{(i)} m_j^{true} \quad \text{with} \quad a_j^{(i)} = \sum_{k=1}^{N} c_k^{(i)} G_{kj} \quad \text{and} \quad c_k^{(i)} = G_{ik}^{-g}$$

relationship to model resolution **R**

$$\mathbf{m}^{\text{est}} = \mathbf{R}\mathbf{m}^{\text{true}} = \mathbf{G}^{-\text{g}}\mathbf{G}\mathbf{m}^{\text{true}} \quad \text{or} \quad m_i^{est} = \sum_{j=1}^{M} R_{ij} m_j^{true} = \sum_{j=1}^{M} \sum_{k=1}^{N} G_{ik}^{-g} G_{kj} m_j^{true}$$

$$m_i^{est} = \langle m \rangle^{(i)} = \sum_{j=1}^{M} a_j^{(i)} m_j^{true} \quad \text{with} \quad a_j^{(i)} = \sum_{k=1}^{N} c_k^{(i)} G_{kj} \quad \text{and} \quad c_k^{(i)} = G_{ik}^{-g}$$

$\mathbf{a}^{\text{T}}$ is a linear combination of the rows of the data kernel $\mathbf{G}$

if we just pick an average
out of the hat
because we like it ... its nicely localized

its not likely that it can be built out of the rows of $\mathbf{G}$

so it will not be unique

suppose we pick a
average that is not unique

is it of any use?

Part 3

bounding localized averages

even though they are nonunique

we will now show

if we can put weak bounds on **m**

they may translate into stronger bounds on <m>

example

$$\mathbf{m}^{\text{gen}} = \begin{bmatrix} d_1 \\ d_1 \\ d_1 \\ d_1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

with

$$\mathbf{a} = ({}^{1}/{}_{3})[1,1,1,0]^{\mathrm{T}}$$

so

$$\langle m \rangle = \mathbf{a}^{\mathrm{T}}\mathbf{m}^{\text{gen}} = d_1 + 0 + 0 + \frac{1}{3}\alpha_3$$

example

$$\mathbf{m}^{\text{gen}} = \begin{bmatrix} d_1 \\ d_1 \\ d_1 \\ d_1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

with

$$\mathbf{a} = ({}^1/{}_3)[1,1,1,0]^{\mathrm{T}}$$

so

$$\langle m \rangle = \mathbf{a}^{\mathrm{T}} \mathbf{m}^{\text{gen}} = d_1 + 0 + 0 + \frac{1}{3}\alpha_3$$

nonunique

but suppose m_i is bounded

$$0 > m_i > 2d_1$$

$$\mathbf{m}^{\text{gen}} = \begin{bmatrix} d_1 \\ d_1 \\ d_1 \\ d_1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

smallest $\alpha_3 = -d_1$

largest $\alpha_3 = +d_1$

smallest $\alpha_3 = -d_1$

largest $\alpha_3 = +d_1$

$$\langle m \rangle = \mathbf{a}^{\mathrm{T}} \mathbf{m}^{\mathrm{gen}} = d_1 + 0 + 0 + \frac{1}{3} \alpha_3$$

$$(2/3)\, d_1 \quad > \quad <m> \quad > \quad (4/3) d_1$$

smallest $\alpha_3 = -d_1$

largest $\alpha_3 = +d_1$

$$\langle m \rangle = \mathbf{a}^{\mathrm{T}}\mathbf{m}^{\mathrm{gen}} = d_1 + 0 + 0 + \frac{1}{3}\alpha_3$$

$$(2/3)\, d_1 \quad > \quad <m> \quad > \quad (4/3) d_1$$

bounds on $<m>$ tighter than bounds on m_i

the question is how to do this in more complicated cases

maximize/minimize $\langle m \rangle = \mathbf{a}^{\mathrm{T}}\mathbf{m}$ with respect to $\mathbf{m}$

with the constraints $\mathbf{Gm} = \mathbf{d}^{\mathrm{obs}}$ and $\mathbf{m}^{(\mathrm{l})} \leq \mathbf{m} \leq \mathbf{m}^{(\mathrm{u})}$

Part 4

The *Linear Programming* Problem

the *Linear Programming* problem

find $\mathbf{x}$ that minimizes $z = \mathbf{f}^{\mathrm{T}}\mathbf{x}$

with the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ and and $\mathbf{C}\mathbf{x} = \mathbf{d}$ and $\mathbf{x}^{(\mathrm{l})} \leq \mathbf{x} \leq \mathbf{x}^{(\mathrm{u})}$

the *Linear Programming* problem

flipping sign switches minimization to maximization

find $\mathbf{x}$ that minimizes $z = \mathbf{f}^{\mathrm{T}}\mathbf{x}$

with the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ and and $\mathbf{C}\mathbf{x} = \mathbf{d}$ and $\mathbf{x}^{(l)} \leq \mathbf{x} \leq \mathbf{x}^{(u)}$

flipping signs of **A** and b switches to $\geq$

in Business

quantity of each product

profit

unit profit

find $\mathbf{x}$ that ~~minimizes~~ $z = \mathbf{f}^{\mathrm{T}}\mathbf{x}$

maximizes

with the constraints $\boxed{\mathbf{A}\mathbf{x} \leq \mathbf{b} \text{ and and } \mathbf{C}\mathbf{x} = \mathbf{d}}$ and $\boxed{\mathbf{x}^{(\mathrm{l})} \leq \mathbf{x}} \leq \mathbf{x}^{(\mathrm{u})}$

physical limitations of factory
government regulations
etc

no negative production

care about both profit z and product quantities **x**

in our case

first minimize
then
maximize

m <m> **a**

find $\mathbf{x}$ that minimizes $z = \mathbf{f}^{\mathrm{T}}\mathbf{x}$

with the constraints $\boxed{\mathbf{Ax} \leq \mathbf{b}}$ and and $\boxed{\mathbf{Cx} = \mathbf{d}}$ and $\mathbf{x}^{(l)} \leq \mathbf{x} \leq \mathbf{x}^{(u)}$

not needed **Gm=d** bounds on **m**

care only about <m>, not **m**

In *MatLab*

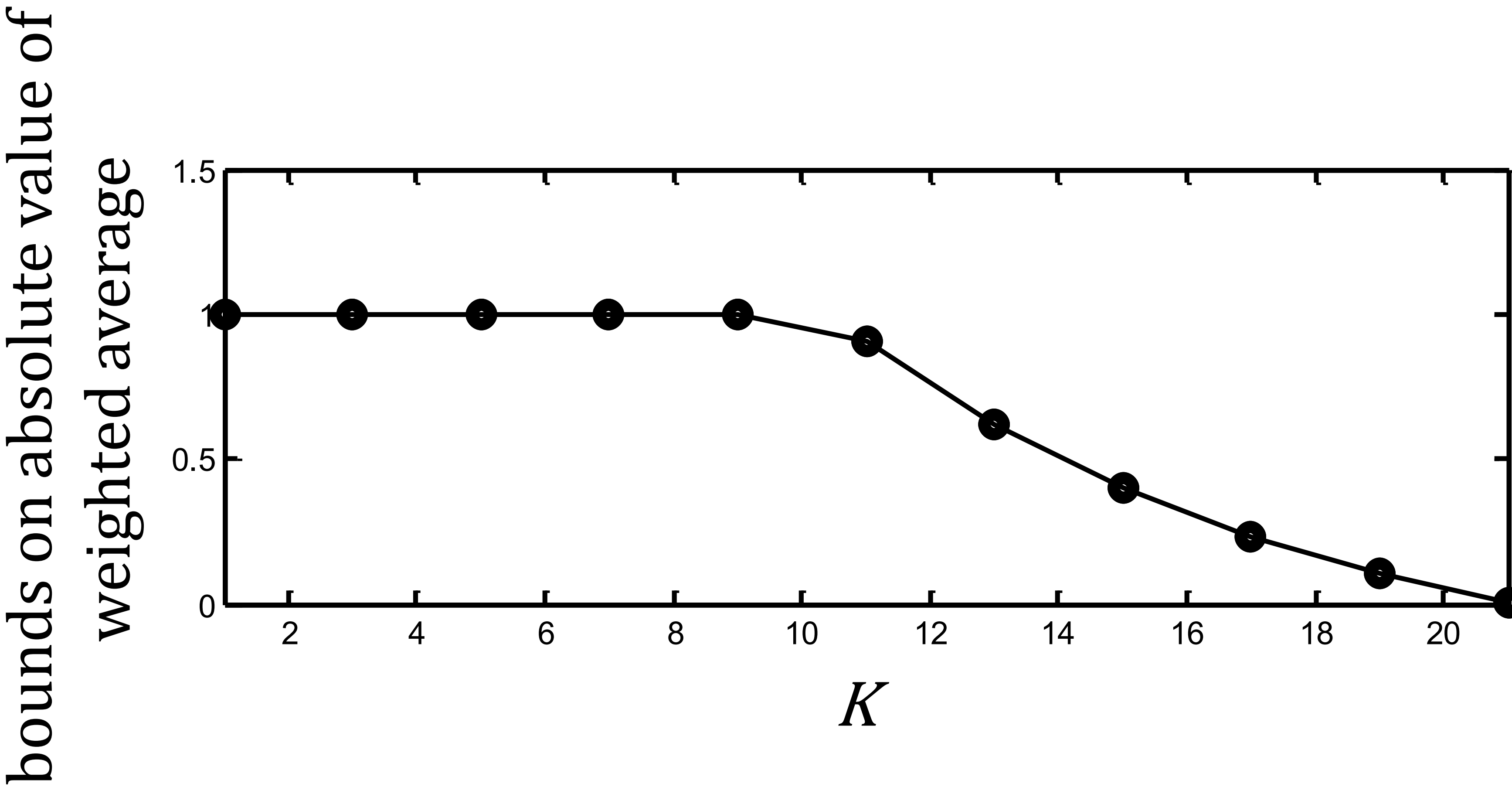
```
[mest1, amin] = linprog(  a,[],[],G,dobs,mlb,mub);
[mest2, amax] = linprog( -a,[],[],G,dobs,mlb,mub);
```

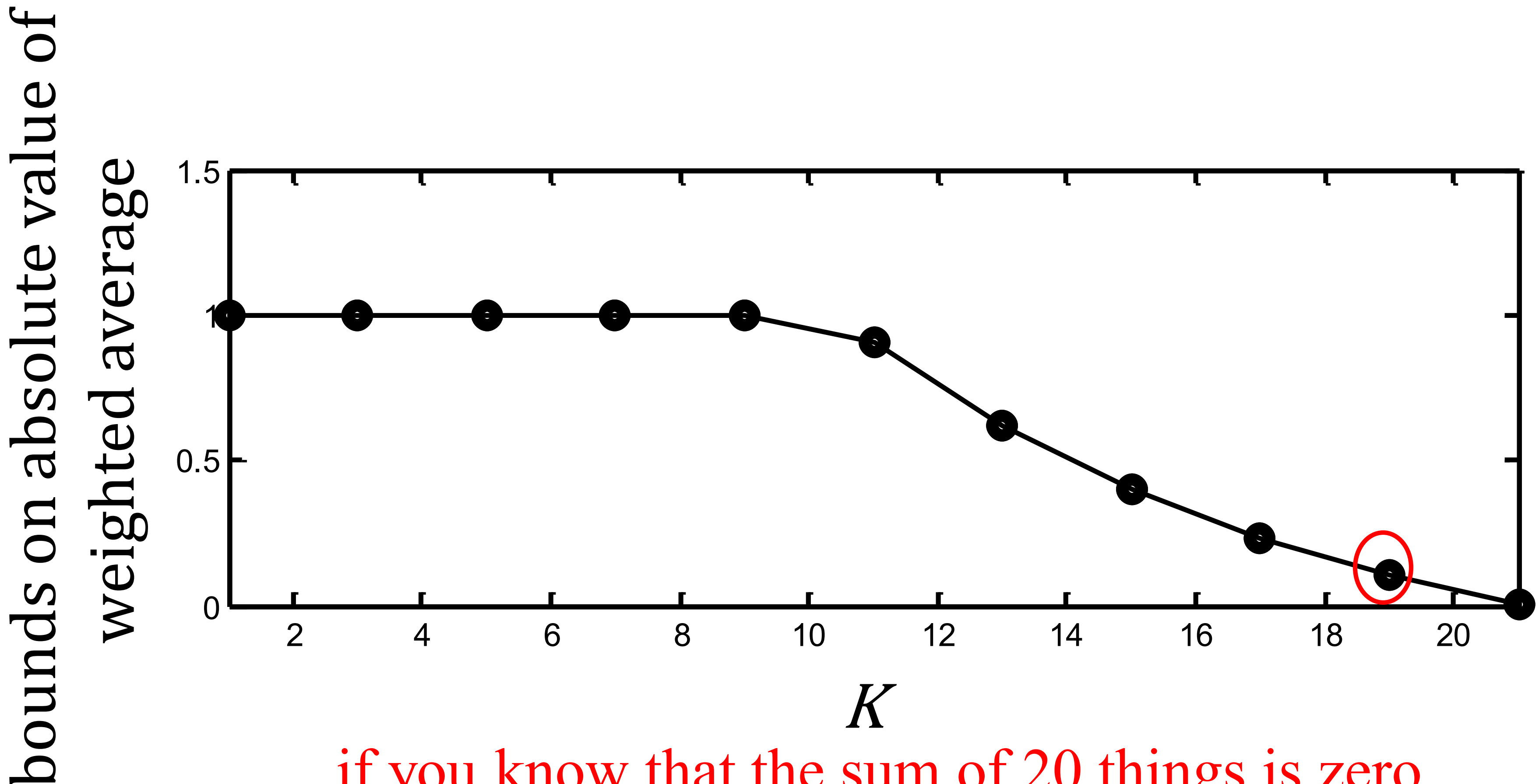
Example 1

simple data kernel
one datum
sum of m_i is zero

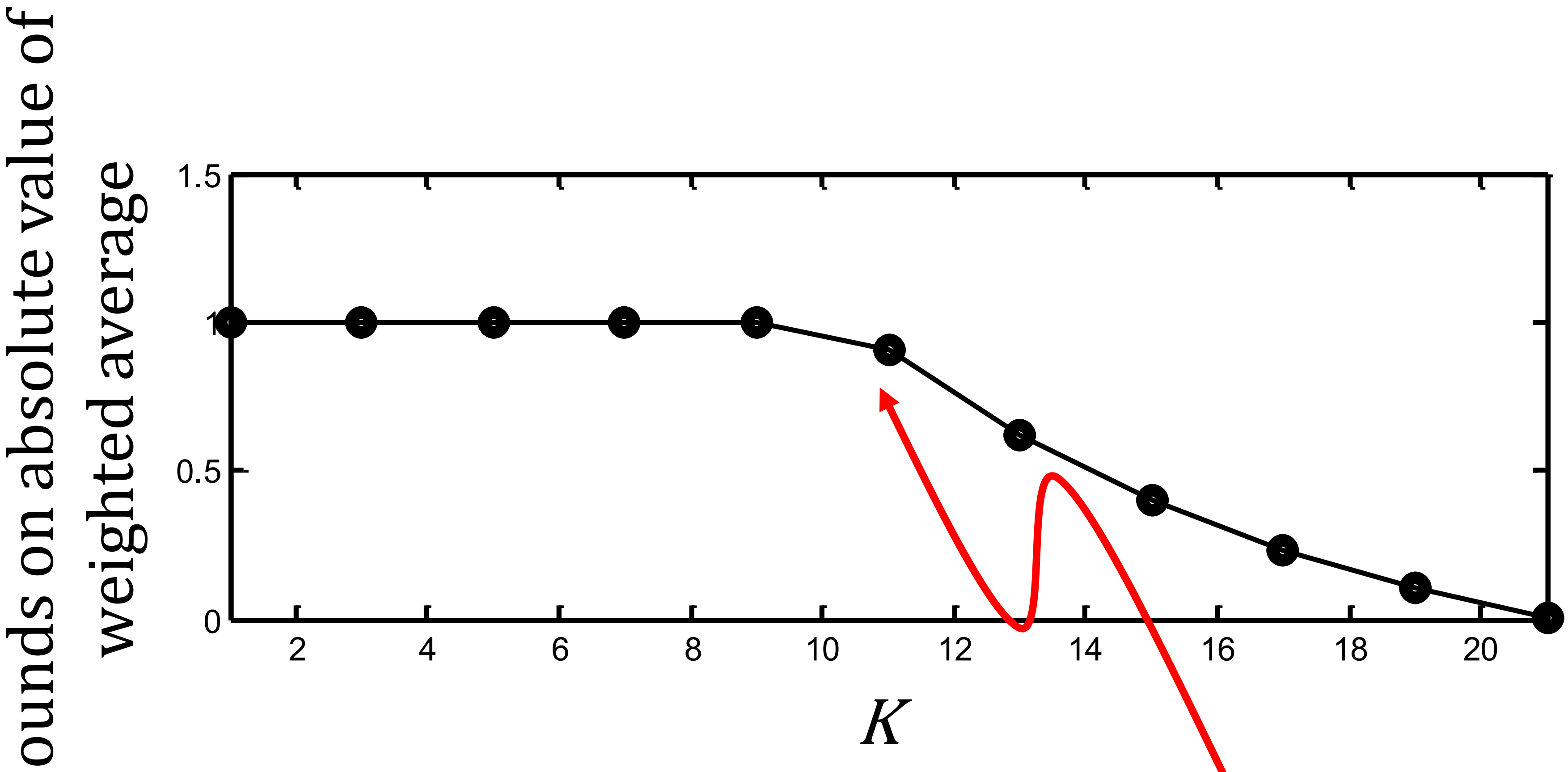
bounds
$|m_i| \leq 1$

average
unweighted average of K model parameters


bounds on absolute value of
weighted average
1.5
1
0.5
0
2
4
6
8
10
12
14
16
18
20
K



if you know that the sum of 20 things is zero
and
if you know that the things are bounded by ± 1
then you know
the sum of 19 of the things is bounded by about ± 0.1



for $K>10$
$<m>$ has tigher bounds
than m_i

Example 2

more complicated data kernel

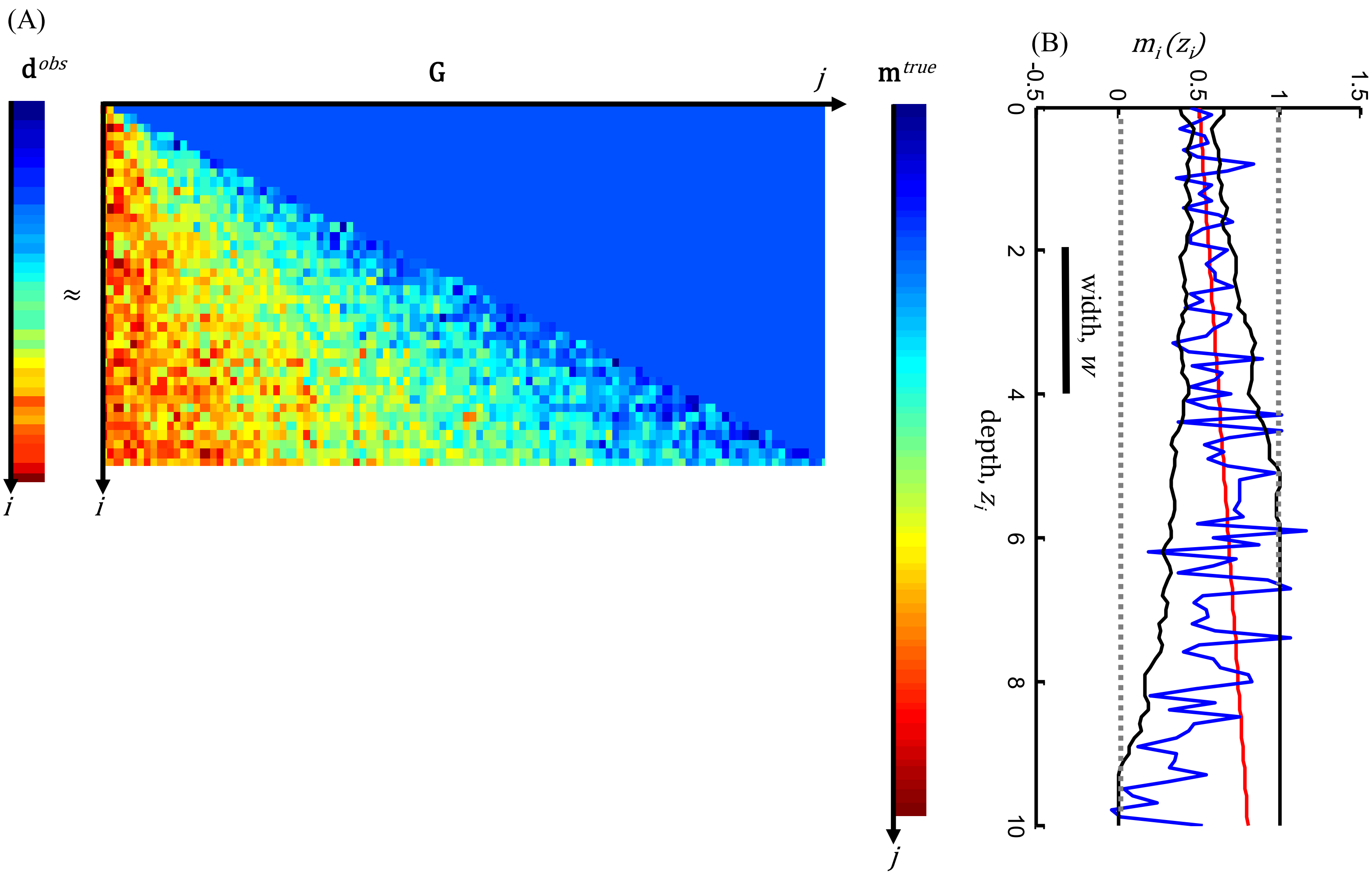
d_k weighted average of first $5k/2$ m's

bounds

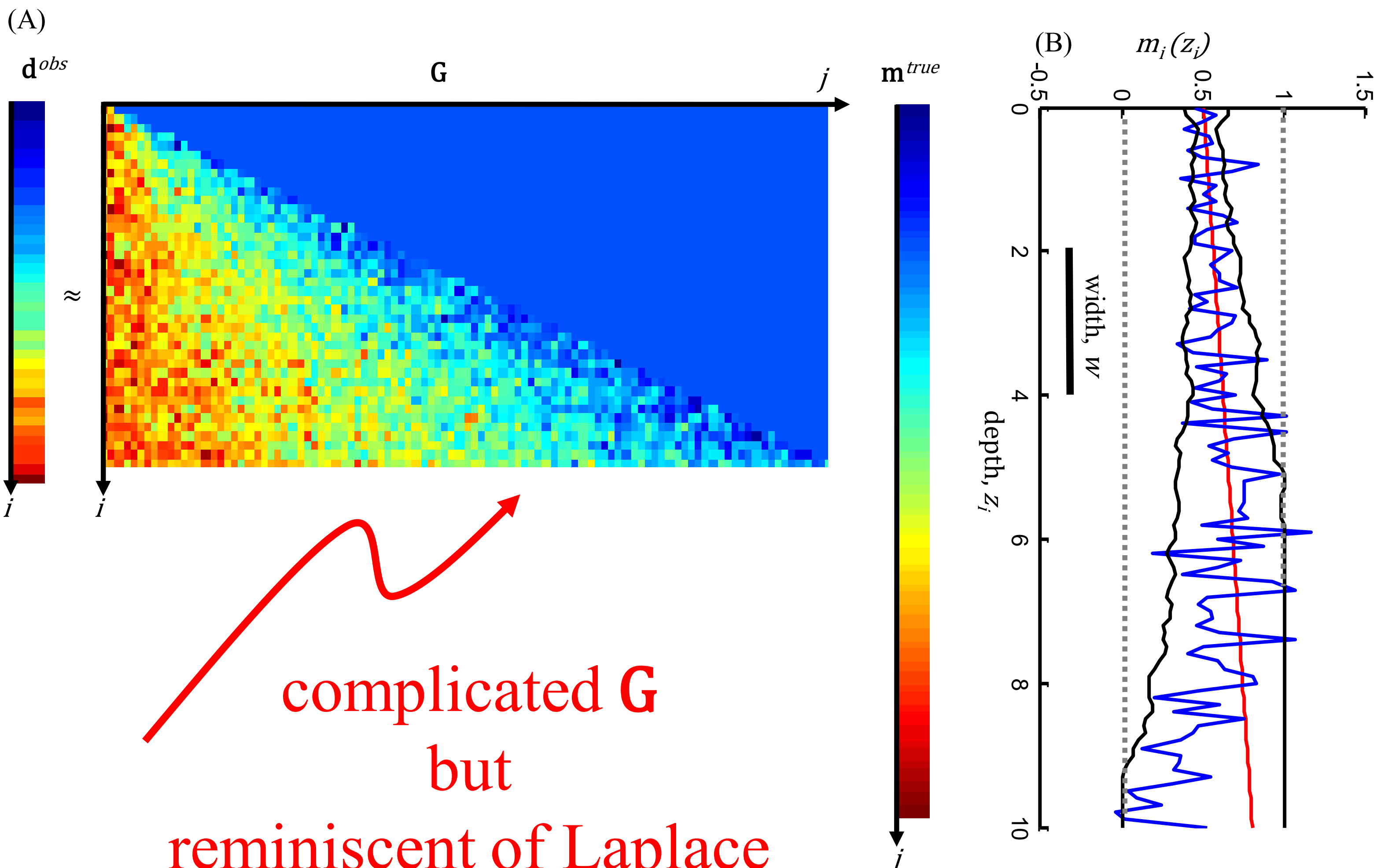
$0 \leq m_i \leq 1$

average

localized average of 5 neighboring model parameters



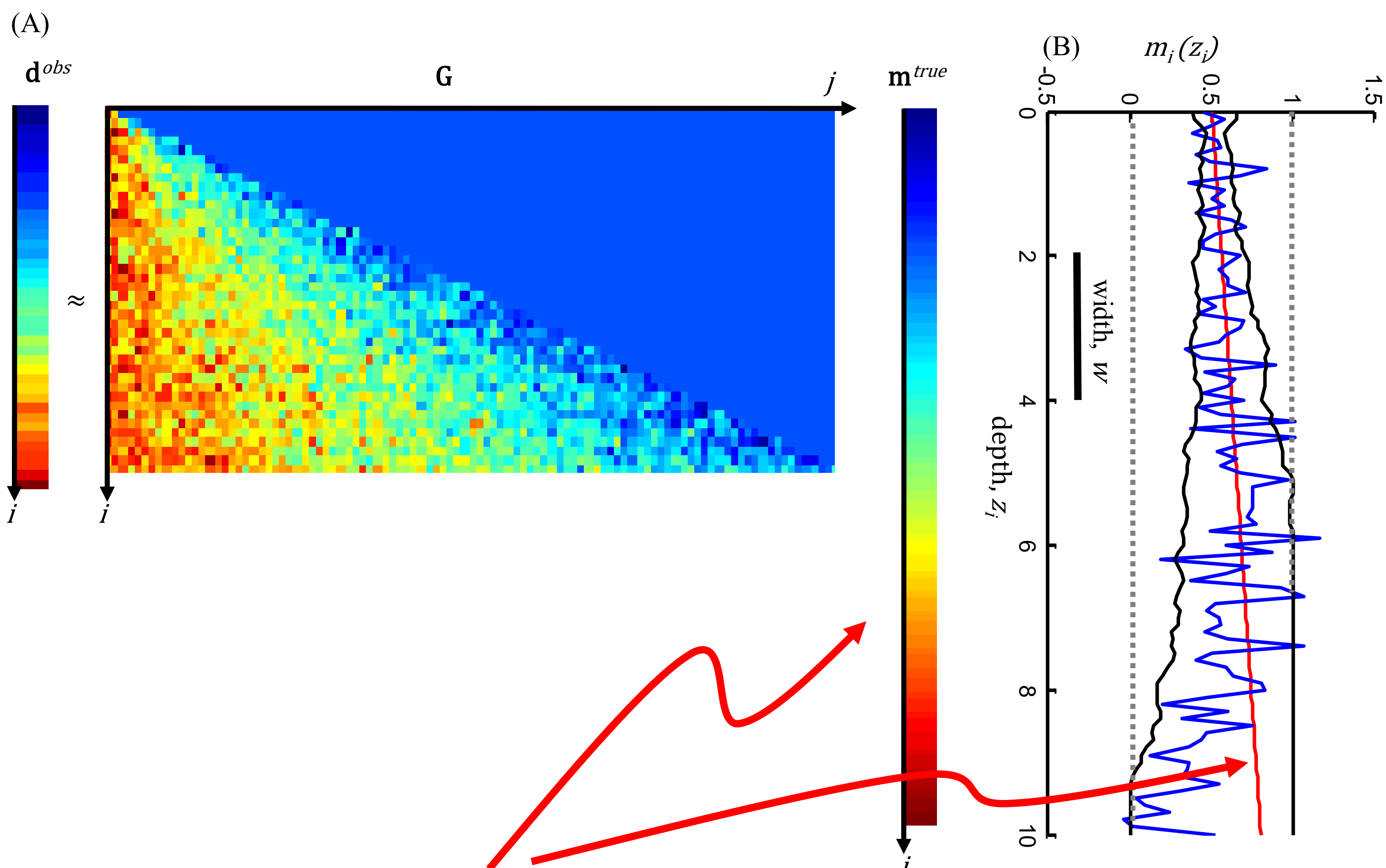
(A)
$\mathbf{d}^{obs}$
$\approx$
G
$\mathbf{m}^{true}$
i
j
(B)
$m_i(z_i)$
-0.5
0
0.5
1
1.5
2
4
6
8
10
width, w
depth, z_i



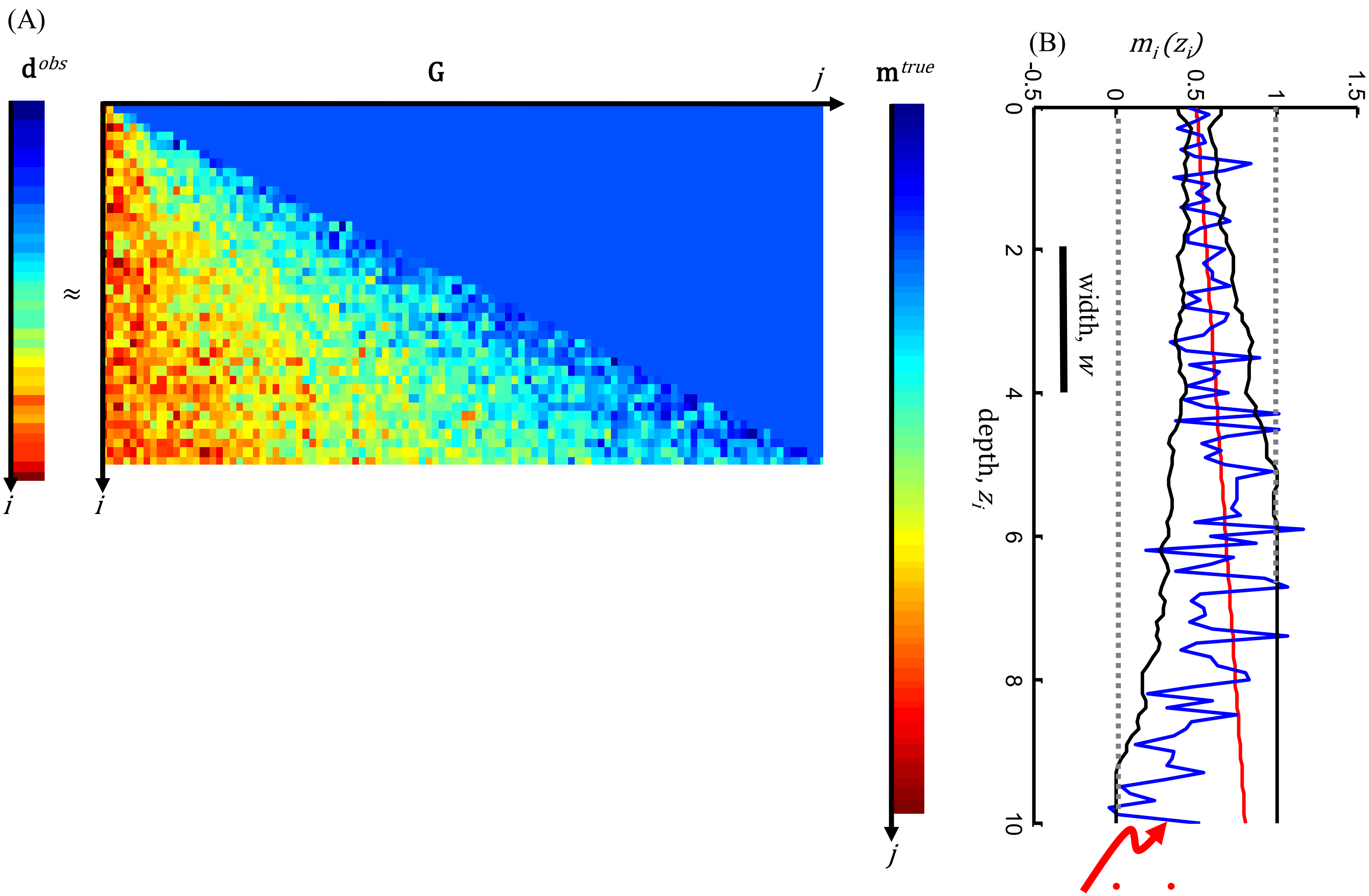
complicated **G**

but

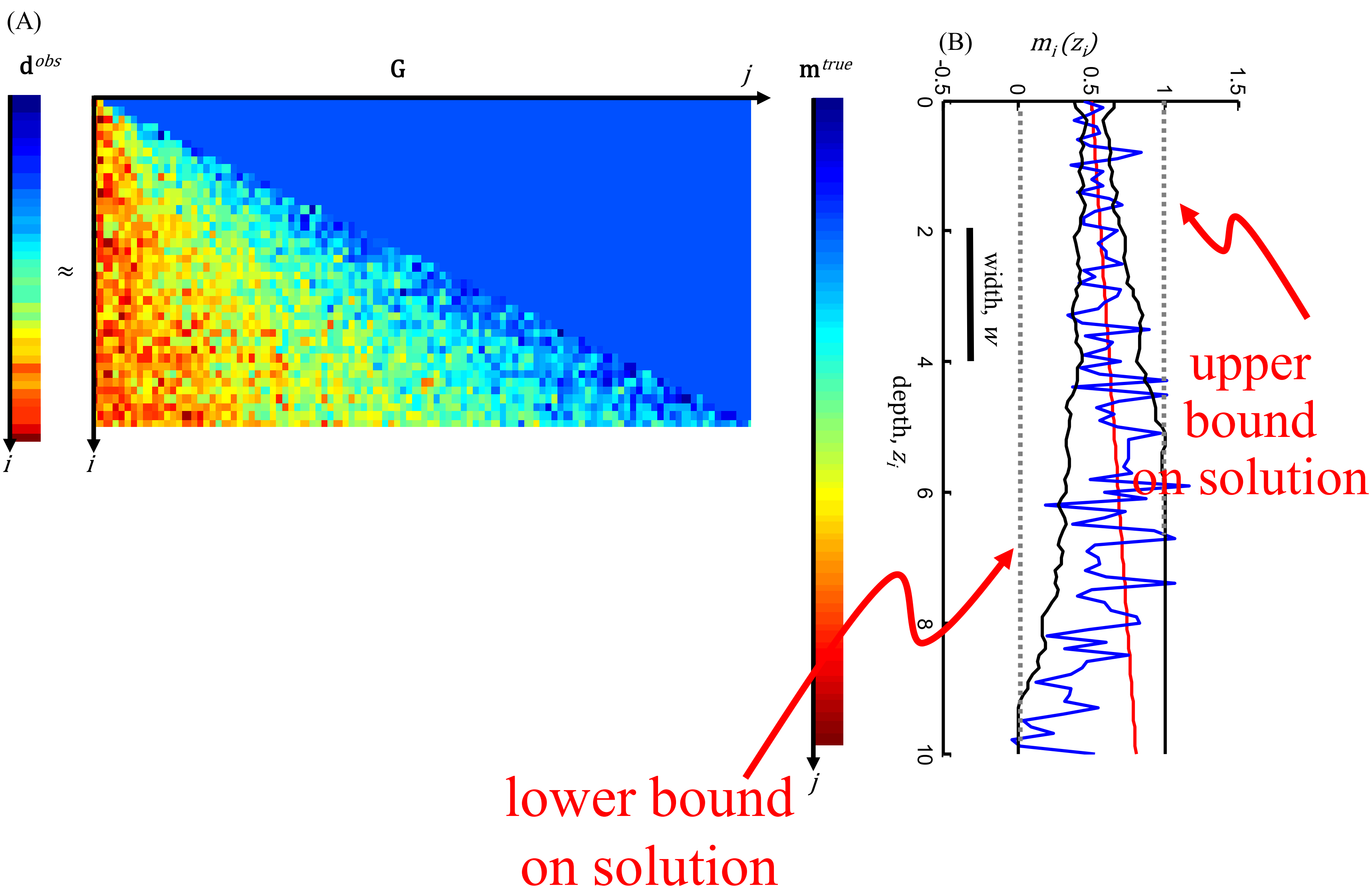
reminiscent of Laplace Transform kernel

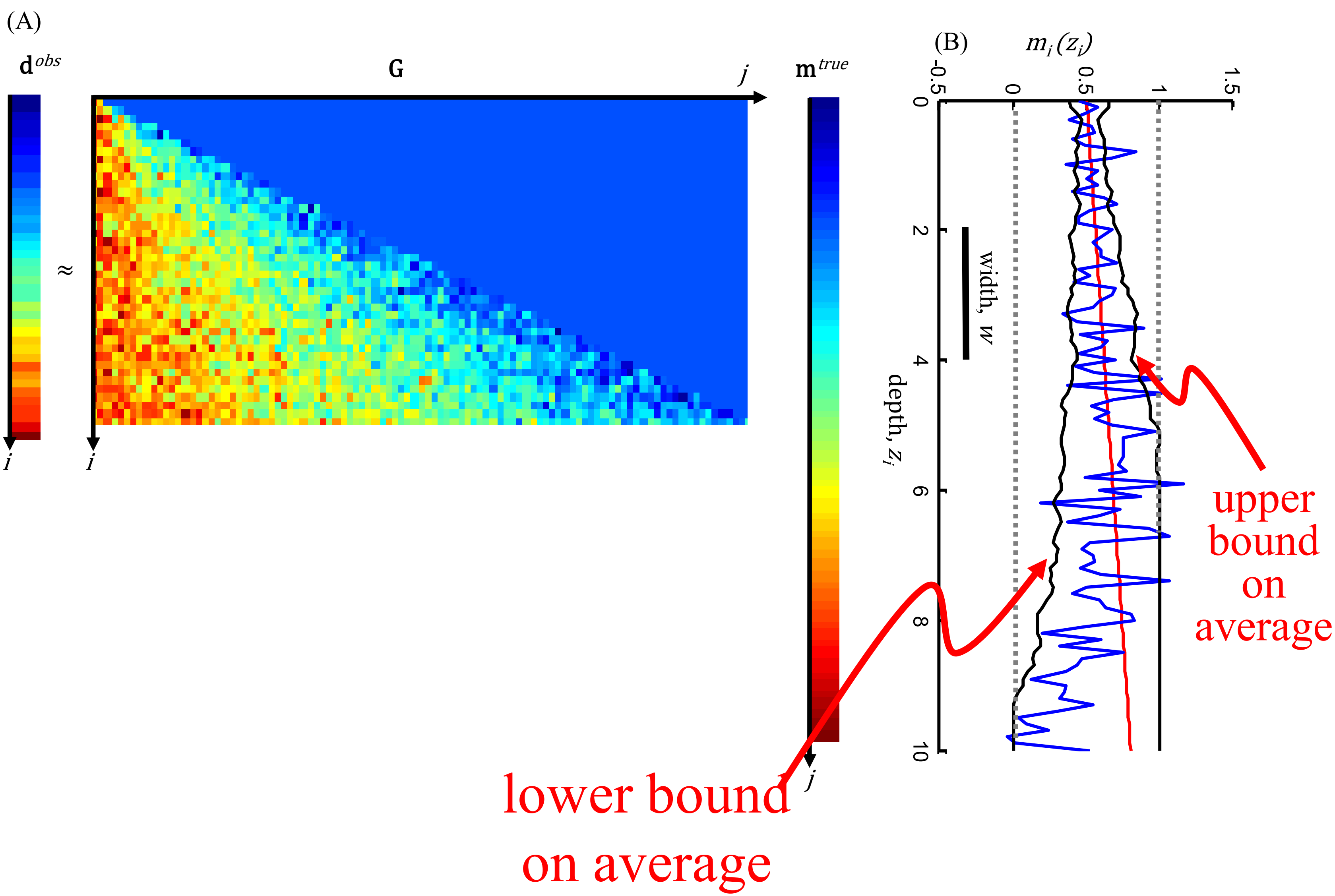


true m_i increased with depth z_i



minimum
length solution


(A)
d^obs
≈
G
j
i
i
m^true
j
(B)
m_i(z_i)
-0.5
0
0.5
1
1.5
0
2
4
6
8
10
width, w
depth, z_i
upper bound on solution
lower bound on solution


(A)
$\mathbf{d}^{obs}$
$\approx$
G
j
$\mathbf{m}^{true}$
i
i
j
(B)
$m_i(z_i)$
-0.5
0
0.5
1
1.5
0
2
4
6
8
10
width, w
depth, z_i
upper bound on average
lower bound on average