Lecture 13

L_1 , L_∞ Norm Problems and Linear Programming

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, Empircal 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

Review Material on Outliers and Long-Tailed Distributions

Derive the L_1 estimate of the
mean and variance of an exponential distribution

Solve the Linear Inverse Problem under the L_1 norm
by Transformation to a Linear Programming Problem

Do the same for the L_∞ problem

Part 1

Review Material on Outliers and Long-Tailed Distributions

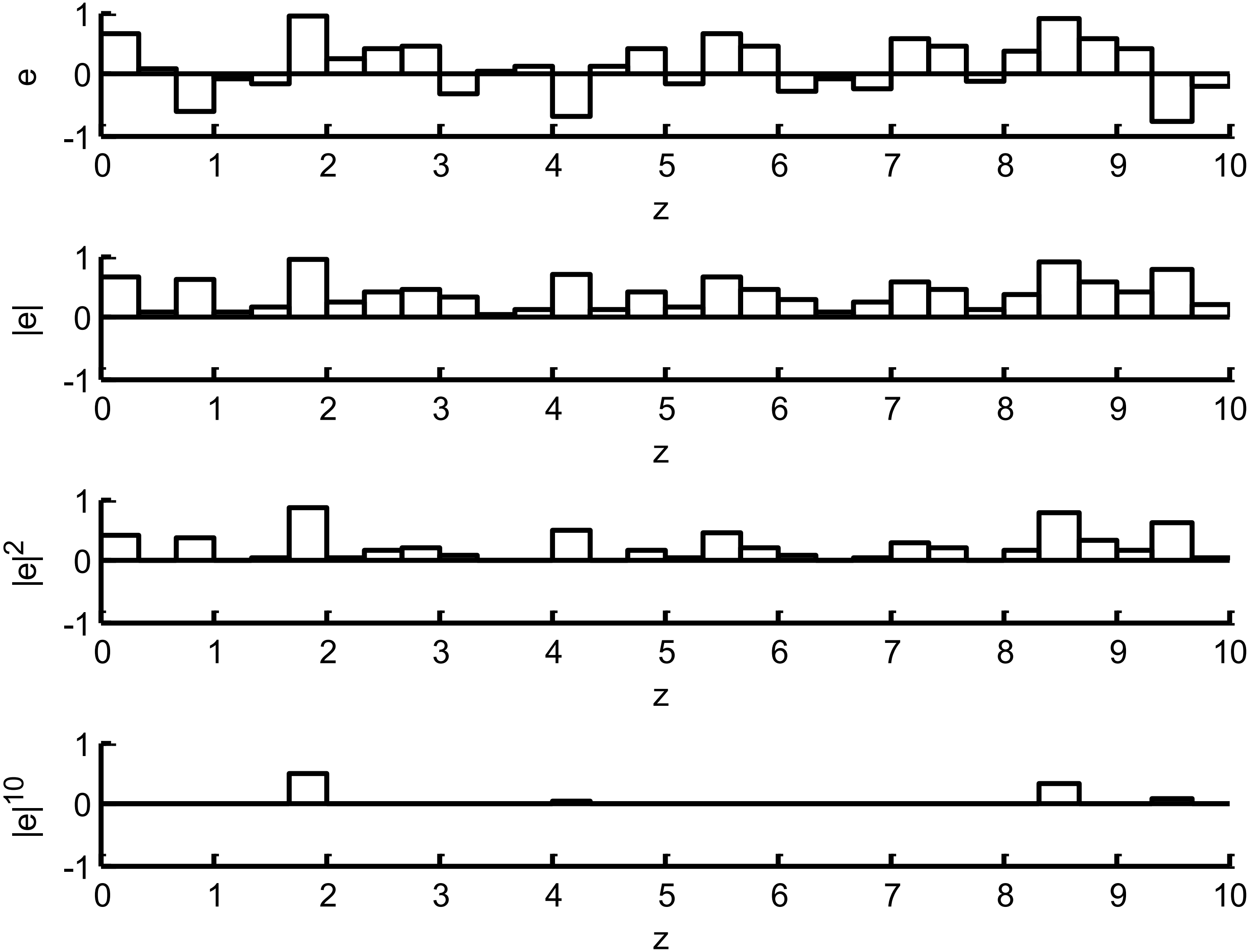
Review of the L_n family of norms

$$L_1 \text{ norm:} \quad \|\mathbf{e}\|_1 = \left[\sum_i |e_i|^1\right]$$

$$L_2 \text{ norm:} \quad \|\mathbf{e}\|_2 = \left[\sum_i |e_i|^2\right]^{1/2}$$

$$L_n \text{ norm:} \quad \|\mathbf{e}\|_n = \left[\sum_i |e_i|^n\right]^{1/n}$$

higher norms give increaing weight to largest element of **e**

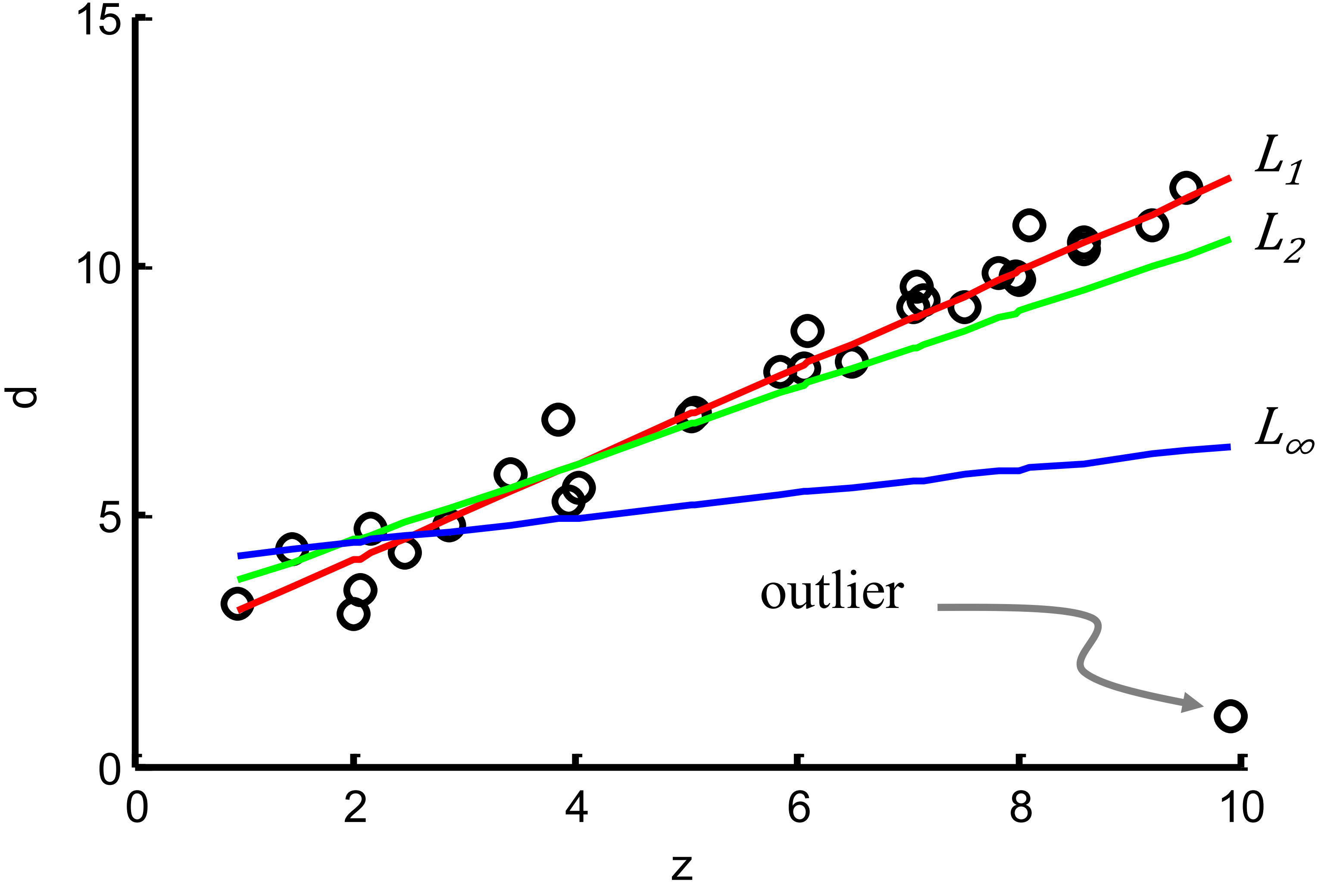


limiting case

$$L_\infty \text{ norm:} \quad \|\mathbf{e}\|_\infty = \max_i |e_i|$$

but which norm to use?

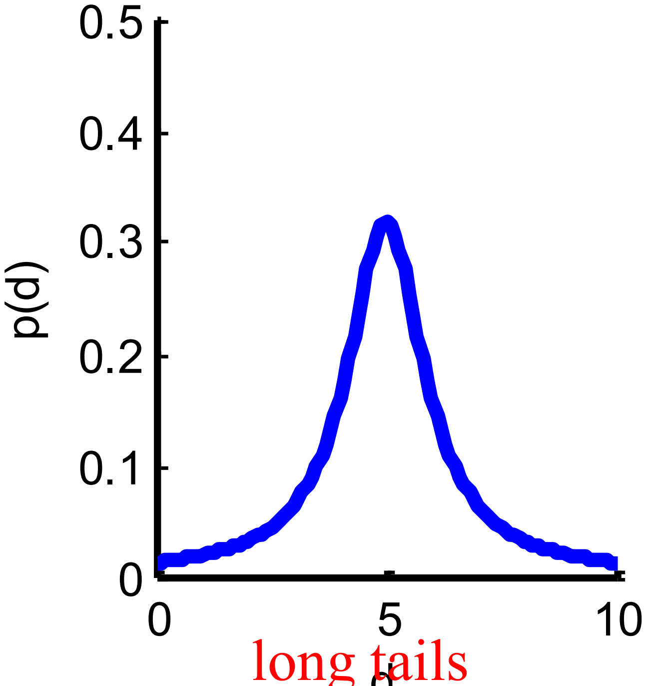
it makes a difference!



15
10
5
0
d
0
2
4
6
8
10
z
L_1
L_2
L_∞
outlier

Answer is related to the distribution of the error. Are outliers common or rare?

A)



long tails

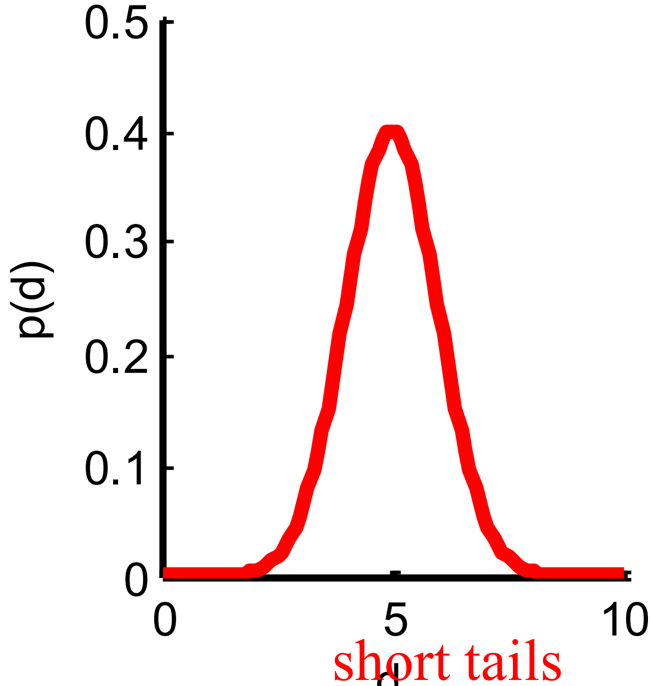
outliers common

outliers unimportant

use low norm

gives low weight to outliers

B)



short tails

outliers uncommon

outliers important

use high norm

gives high weight to outliers

as we showed previously ...

use L_2 norm
when data has
Gaussian-distributed error

as we will show in a moment ...

use L_1 norm
when data has
Exponentially-distributed error

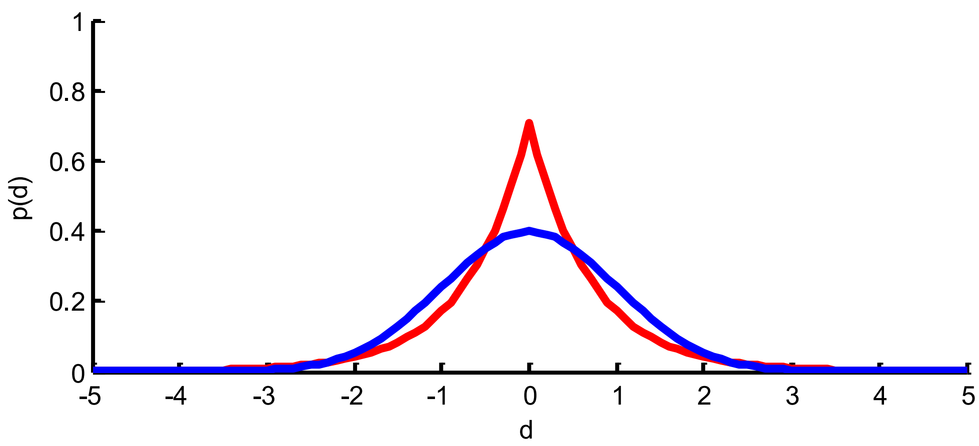
comparison of p.d.f.'s

Gaussian

$$p(d) = \frac{1}{(2\pi)^{½}\sigma} \exp\left[-\frac{(d-\langle d \rangle)^2}{2\sigma^2}\right]$$

Exponential

$$p(d) = \frac{1}{(2)^{½}\sigma} \exp\left\{-(2)^{½}\frac{|d-\langle d \rangle|}{\sigma}\right\}$$



to make realizations of an exponentially-distributed random variable in MatLab

```
mu = sd/sqrt(2);
rsign = (2*(random('unid',2,Nr,1)-1)-1);
dr = dbar + rsign .* ...
          random('exponential',mu,Nr,1);
```

Part 2

Derive the L_1 estimate of the mean and variance of an exponential distribution

use of Principle of Maximum Likelihood

maximize

$$L = \log p(\mathbf{d}^{\text{obs}})$$

the log-probability that the observed data was in fact observed

with respect to unknown parameters in the p.d.f.

e.g. its mean m_1 and variance σ^2

Previous Example: Gaussian p.d.f.

$$p(\mathbf{d}) = \sigma^{-N}(2\pi)^{-N/2}\exp\left[-½\sigma^{-2}\sum_{i=1}^{N}[d_i - m_1]^2\right]$$

$$L = \log(p(\mathbf{d}^{obs})) = -N log(\sigma) - ½\sigma^{-2}\sum_{i=1}^{N}\left(d_i^{obs} - m_1\right)^2$$

$$\frac{\partial L}{\partial m_1} = 0 = ½\sigma^{-2} 2m_1 \sum_{i=1}^{N}\left(d_i^{obs} - m_1\right)$$

$$\frac{\partial L}{\partial \sigma} = 0 = -\frac{N}{\sigma} + \sigma^{-3}\sum_{i=1}^{N}\left(d_i^{obs} - m_1\right)^2$$

solving the two equations

$$m_1^{est} = \frac{1}{N} \sum_{i=1}^{N} d_i^{obs} \quad \text{and} \quad \sigma^{est} = \left[\frac{1}{N} \sum_{i=1}^{N} \left(d_i^{obs} - m_1 \right)^2 \right]^{1/2}$$

solving the two equations

$$m_1^{est} = \frac{1}{N}\sum_{i=1}^{N} d_i^{obs} \quad \text{and} \quad \sigma^{est} = \left[\frac{1}{N}\sum_{i=1}^{N}\left(d_i^{obs} - m_1\right)^2\right]^{½}$$

usual formula for the sample mean

almost the usual formula for the sample standard deviation

New Example: Exponential p.d.f.

$$p(\mathbf{d}) = (2)^{-N/2}\sigma^{-N}\exp\left[-\frac{(2)^{½}}{\sigma}\sum_{i=1}^{N}|d_i - m_1|\right]$$

$$\text{maximize} \quad L = \log P = -\frac{N}{2}\log(2) - N\log(\sigma) - \frac{(2)^{½}}{\sigma}\sum_{i=1}^{N}|d_i - m_1|$$

$$\frac{\partial L}{\partial m_1} = 0 = -\frac{(2)^{½}}{\sigma}\sum_{i=1}^{N}\text{sign}(d_i - m_1) \quad \text{and} \quad \frac{\partial L}{\partial \sigma} = 0 = \frac{N}{\sigma} - \frac{(2)^{½}}{\sigma^2}\sum_{i=1}^{N}|d_i - m_1|$$

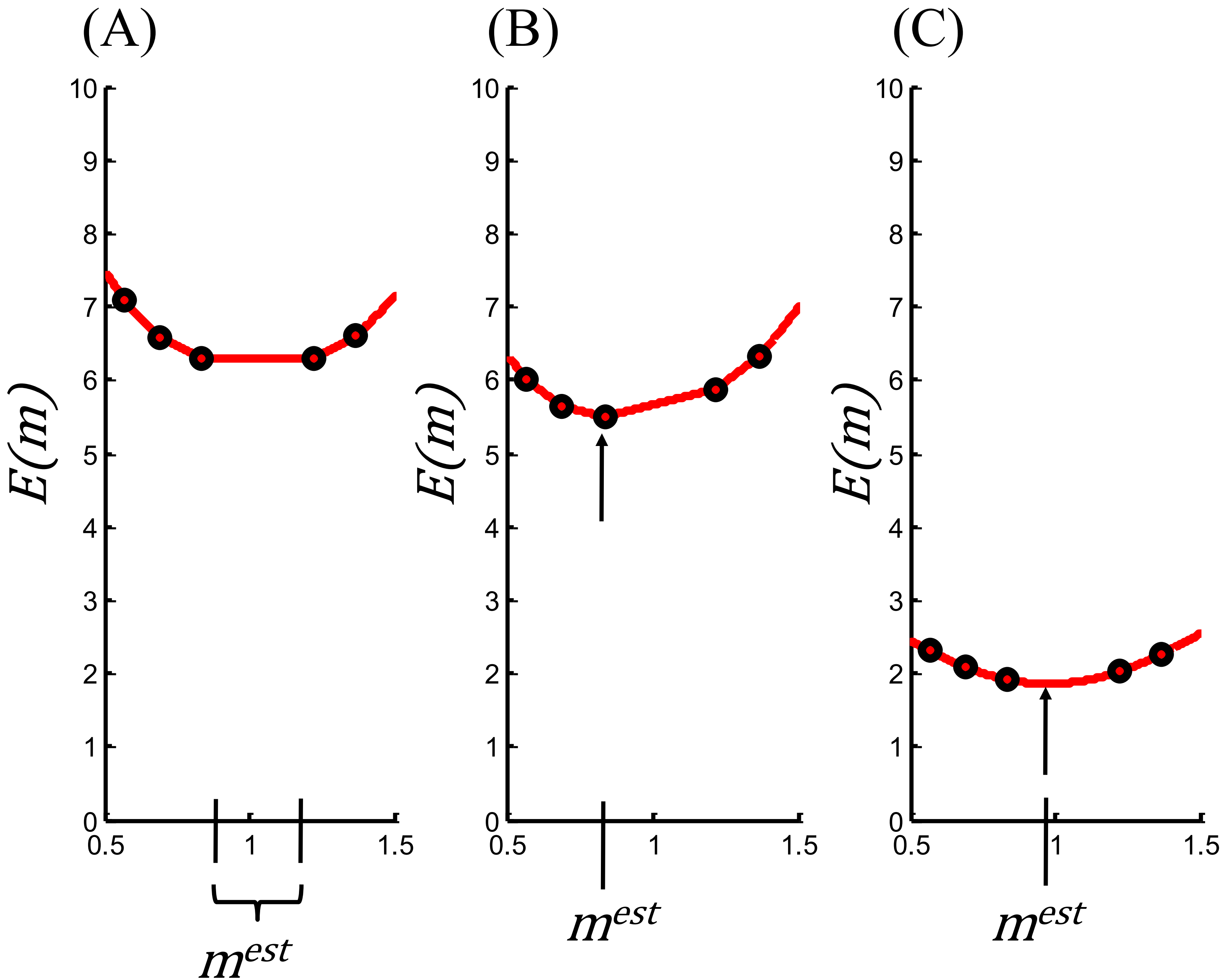
solving the two equations

m_1^{est}=median($\mathbf{d}$) and $\sigma^{est} = \frac{(2)^{½}}{N} \sum_{i=1}^{N} |d_i - m_1|$

solving the two equations

$$m_1^{est} = \text{median}(\mathbf{d}) \quad \text{and} \quad \sigma^{est} = \frac{(2)^{1/2}}{N} \sum_{i=1}^{N} |d_i - m_1|$$

more robust than sample mean since outlier moves it only by "one data point"



(A)
(B)
(C)
E(m)
0
1
2
3
4
5
6
7
8
9
10
0.5
1
1.5
m^est

observations

1. When the number of data are even, the solution in non-unique but bounded
2. The solution exactly satisfies one of the data

these properties carry over to the general linear problem

1. In certain cases, the solution can be non-unique but bounded
2. The solution exactly satisfies M of the data equations

Part 3

Solve the Linear Inverse Problem under the L_1 norm by Transformation to a Linear Programming Problem

review

the *Linear Programming* problem

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

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

Case A

The Minimum L_1 Length Solution

minimize

$$L = \sum_{i=1}^{M} \frac{|m_i - \langle m_i \rangle|}{\sigma_{mi}}$$

subject to the constraint

$$\mathbf{Gm} = \mathbf{d}$$

minimize

$$L = \sum_{i=1}^{M} \frac{|m_i - \langle m_i \rangle|}{\sigma_{mi}}$$

weighted L_1
solution length
(weighted by σ_m^{-1})

subject to the constraint

$$\mathbf{Gm}=\mathbf{d}$$

usual data
equations

transformation to an equivalent linear programming problem

minimize $\quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad$ subject to the constraints

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

all variables are required to be positive

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

usual data equations with **m=m'-m''**

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x}_\mathbf{i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

“slack variables”

standard trick in linear programming

to allow **m** to have any sign while $\mathbf{m_1}$ and $\mathbf{m_2}$ are non-negative

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

same as

$$\boldsymbol{\alpha} - \mathbf{x} = [\mathbf{m} - \langle \mathbf{m} \rangle] \quad \text{and} \quad \boldsymbol{\alpha} - \mathbf{x}' = -[\mathbf{m} - \langle \mathbf{m} \rangle]$$

minimize $\quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}}$ subject to the constraints

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

if +

$$\boldsymbol{\alpha} - \mathbf{x} = [\mathbf{m} - \langle \mathbf{m} \rangle] \quad \text{and} \quad \boldsymbol{\alpha} - \mathbf{x}' = -[\mathbf{m} - \langle \mathbf{m} \rangle]$$

then $\boldsymbol{\alpha} \geq (\mathbf{m}\text{-}\langle\mathbf{m}\rangle)$
since $\mathbf{x} \geq 0$

can always be satisfied by
choosing an appropriate **x**’

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

if -

$$\boldsymbol{\alpha} - \mathbf{x} = [\mathbf{m} - \langle \mathbf{m} \rangle] \quad \text{and} \quad \boldsymbol{\alpha} - \mathbf{x}' = -[\mathbf{m} - \langle \mathbf{m} \rangle]$$

can always be satisfied by choosing an appropriate **x**'

then $\boldsymbol{\alpha} \geq -(\mathbf{m} - \langle \mathbf{m} \rangle)$ since $\mathbf{x} \geq 0$

$$\text{minimize} \quad z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}} \quad \text{subject to the constraints}$$

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle \mathbf{m} \rangle$$

$$\boldsymbol{\alpha} - \mathbf{x} = [\mathbf{m} - \langle \mathbf{m} \rangle] \quad \text{and} \quad \boldsymbol{\alpha} - \mathbf{x}' = -[\mathbf{m} - \langle \mathbf{m} \rangle]$$

taken together then $\boldsymbol{\alpha} \geq |\mathbf{m} - \langle \mathbf{m} \rangle|$

minimize $z = \sum_{i=1}^{M} \frac{\alpha_i}{\sigma_{mi}}$ subject to the constraints

minimizing **z**
same as minimizing
weighted solution length

$$\mathbf{G}(\mathbf{m}' - \mathbf{m}'') = \mathbf{d} \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' + \mathbf{x_i} - \boldsymbol{\alpha} = \langle\mathbf{m}\rangle \quad \text{and} \quad \mathbf{m}' - \mathbf{m}'' - \mathbf{x}' - \boldsymbol{\alpha} = \langle\mathbf{m}\rangle$$

$$\boldsymbol{\alpha} - \mathbf{x} = [\mathbf{m} - \langle\mathbf{m}\rangle] \quad \text{and} \quad \boldsymbol{\alpha} - \mathbf{x}' = -[\mathbf{m} - \langle\mathbf{m}\rangle]$$

Case B

Least L_1 error solution (analogous to least squares)

transformation to an equivalent linear programming problem

$$\text{minimize} \quad E = \sum_{i=1}^{N} \frac{\alpha_i}{\sigma_{di}} \quad \text{subject to the constraints}$$

$$\mathbf{G}[\mathbf{m}' - \mathbf{m}''] + \mathbf{x} - \boldsymbol{\alpha} = \mathbf{d} \quad \text{and} \quad \mathbf{G}[\mathbf{m}' - \mathbf{m}''] - \mathbf{x}' + \boldsymbol{\alpha} = \mathbf{d}$$

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

$$\text{minimize} \quad E = \sum_{i=1}^{N} \frac{\alpha_i}{\sigma_{di}} \quad \text{subject to the constraints}$$

$$\mathbf{G}[\mathbf{m}' - \mathbf{m}''] + \mathbf{x} - \boldsymbol{\alpha} = \mathbf{d} \quad \text{and} \quad \mathbf{G}[\mathbf{m}' - \mathbf{m}''] - \mathbf{x}' + \boldsymbol{\alpha} = \mathbf{d}$$

and

$$\mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \boldsymbol{\alpha} \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

same as

$$\boldsymbol{\alpha} - \mathbf{x} = \mathbf{Gm} - \mathbf{d}$$

$$\boldsymbol{\alpha} - \mathbf{x}' = -(\mathbf{Gm} - \mathbf{d})$$

so previous argument applies

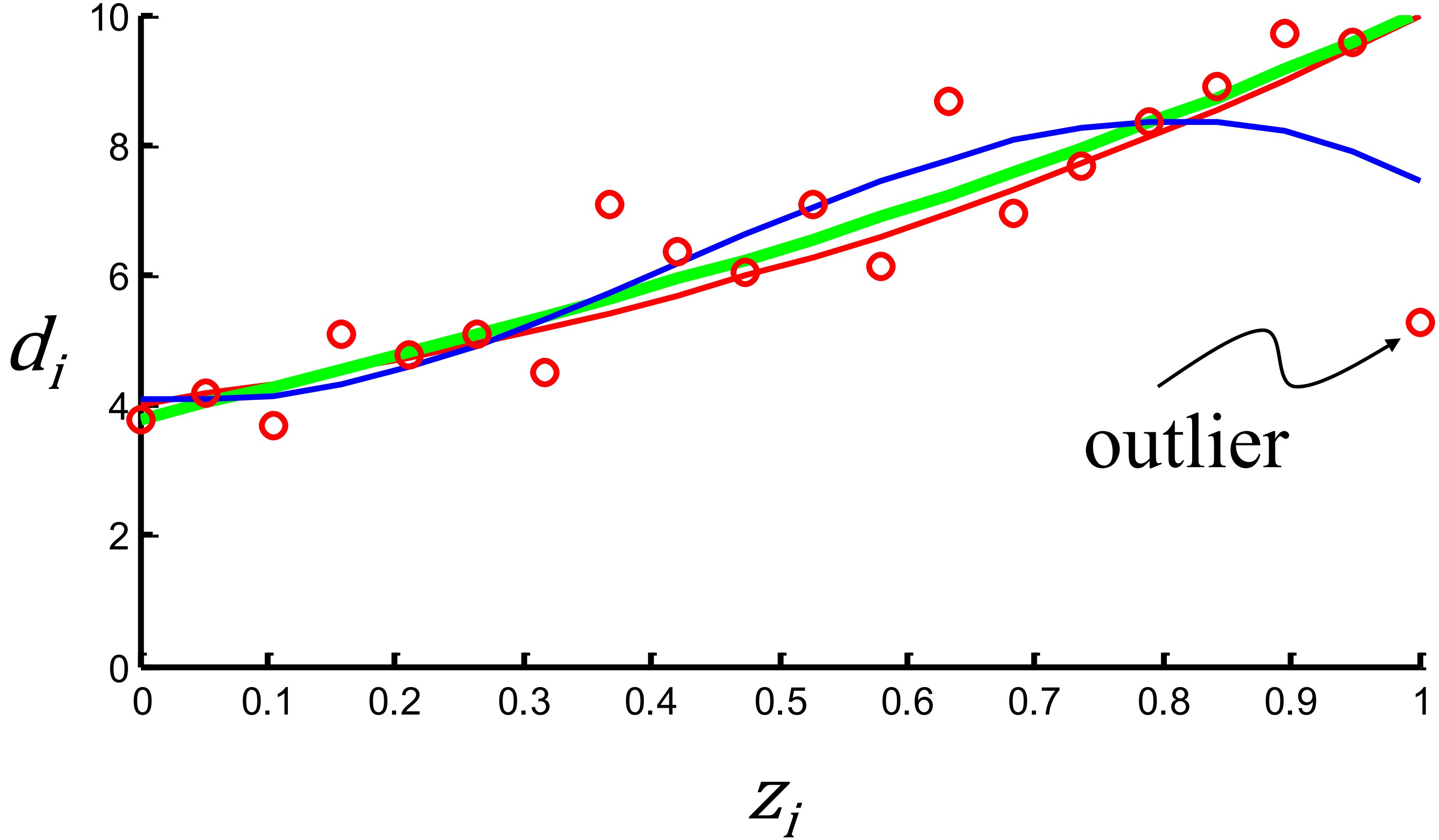
MatLab

```
% variables
% m = mp - mpp
% x = [mp', mpp', alpha', x', xp']'
% mp, mpp len M and alpha, x, xp, len N
L = 2*M+3*N;
x = zeros(L,1);
f = zeros(L,1);
f(2*M+1:2*M+N)=1./sd;
```

```
% equality constraints
Aeq = zeros(2*N,L);
beq = zeros(2*N,1);

% first equation G(mp-mpp)+x-alpha=d
Aeq(1:N,1:M)                    =  G;
Aeq(1:N,M+1:2*M)                = -G;
Aeq(1:N,2*M+1:2*M+N)            = -eye(N,N);
Aeq(1:N,2*M+N+1:2*M+2*N)        =  eye(N,N);
beq(1:N)                        =  dobs;

% second equation G(mp-mpp)-xp+alpha=d
Aeq(N+1:2*N,1:M)                      =  G;
Aeq(N+1:2*N,M+1:2*M)                  = -G;
Aeq(N+1:2*N,2*M+1:2*M+N)              =  eye(N,N);
Aeq(N+1:2*N,2*M+2*N+1:2*M+3*N)        = -eye(N,N);
beq(N+1:2*N)                          =  dobs;
```

```
% inequality constraints A x <= b

% part 1: everything positive
A = zeros(L+2*M,L);
b = zeros(L+2*M,1);
A(1:L,:) = -eye(L,L);
b(1:L) = zeros(L,1);

% part 2; mp and mpp have an upper bound.
A(L+1:L+2*M,:) = eye(2*M,L);
mls = (G'*G)\(G'*dobs); % L2
mupperbound=10*max(abs(mls));
b(L+1:L+2*M) = mupperbound;
```

```
% solve linear programming problem
[x, fmin] = linprog(f,A,b,Aeq,beq);
fmin=-fmin;
mest = x(1:M) - x(M+1:2*M);
```



d_i
z_i
outlier
0
2
4
6
8
10
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

the mixed-determined problem of

minimizing $L+E$

can also be solved via transformation

but we omit it here

Part 4

Solve the Linear Inverse Problem under the L_∞ norm by Transformation to a Linear Programming Problem

we're going to skip all the details

and just show the transformation for the overdetermined case minimize $E = \max_i (e_i / \sigma_{di})$ where $\mathbf{e} = \mathbf{d}^{obs} - \mathbf{Gm}$

minimize α subject to the constraints

$$\sum_{j=1}^{M} G_{ij}[m'_j - m''_j] + x_i - \alpha\sigma_{di} = d_i \quad \text{and} \quad \sum_{j=1}^{M} G_{ij}[m'_j - m''_j] - x'_i + \alpha\sigma_{di} = d_i$$

and $\mathbf{m}' \geq 0$ and $\mathbf{m}'' \geq 0$ and $\alpha \geq 0$ and $\mathbf{x} \geq 0$ and $\mathbf{x}' \geq 0$

note **α** is a scalar

minimize α subject to the constraints

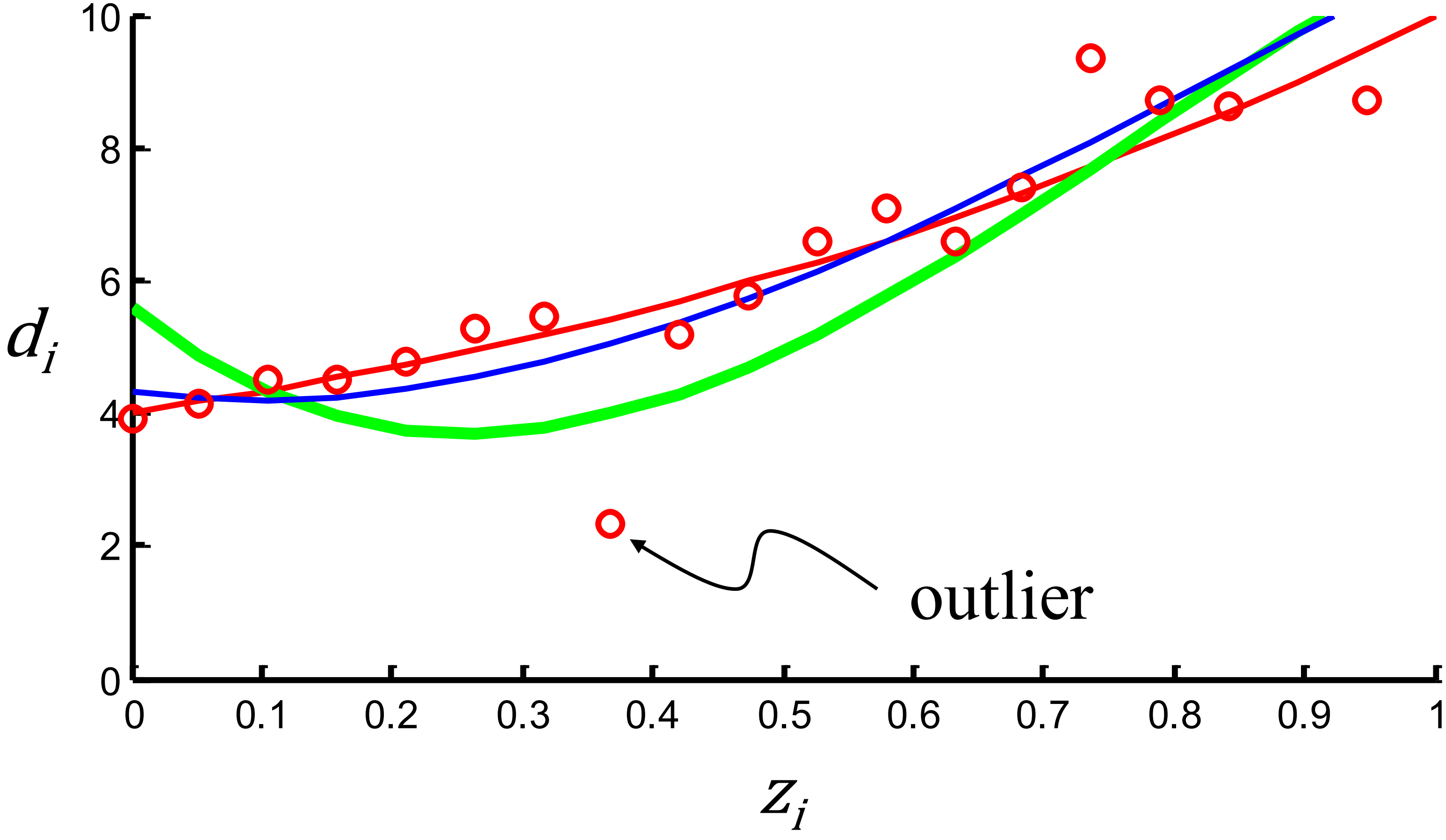
$$\sum_{j=1}^{M} G_{ij}\left[m'_j - m''_j\right] + x_i - \alpha\sigma_{di} = d_i \quad \text{and} \quad \sum_{j=1}^{M} G_{ij}\left[m'_j - m''_j\right] - x'_i + \alpha\sigma_{di} = d_i$$

$$\text{and} \quad \mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \alpha \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$

minimize α subject to the constraints

$$\sum_{j=1}^{M} G_{ij}\left[m'_j - m''_j\right] + x_i - \alpha\sigma_{di} = d_i \quad \text{and} \quad \sum_{j=1}^{M} G_{ij}\left[m'_j - m''_j\right] - x'_i + \alpha\sigma_{di} = d_i$$

$$\text{and} \quad \mathbf{m}' \geq 0 \quad \text{and} \quad \mathbf{m}'' \geq 0 \quad \text{and} \quad \alpha \geq 0 \quad \text{and} \quad \mathbf{x} \geq 0 \quad \text{and} \quad \mathbf{x}' \geq 0$$


10
8
6
4
2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
d_i
z_i
outlier