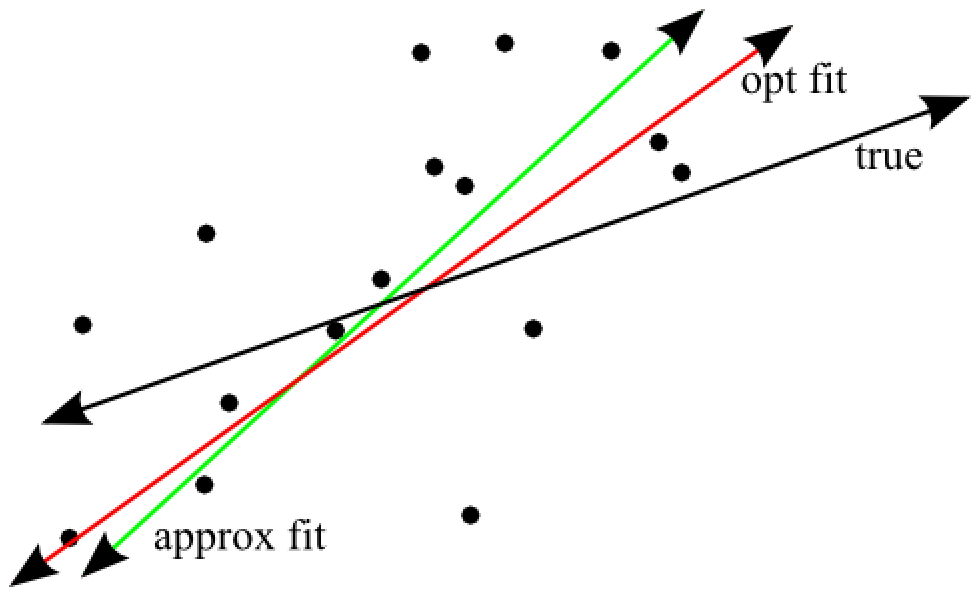


Approximation and fitting

Sean Woodham, Charly Stephens-Cohen, Fei Yang
and Maxim Zalutskiy

Outline

- Norm approximation and least-norm problems
- Regularized approximation
- Robust approximation
- Function fitting and interpolation
- Matlab Examples

Norm Approximation

- Simplest form of a **Norm Approximation** problem is , *minimize* $||Ax - b||$
- $A \in \boldsymbol{R}^{m \times n}$ and $b \in \boldsymbol{R}^m$ are the *problem data*, $x \in \boldsymbol{R}^n$ is the *variable*
- Refer to the solution of the problem as an approximation solution $Ax \approx b$, where the vector $r = Ax - b$ is the *residual*.

Properties of the Norm Approximation Problem

- It is always *convex* and *solvable*.
- Its *optimal value* is zero if and only if $b \in R(A)$
- For the cases where b not in $R(A)$, the columns of A are *linearly independent* and in the matrix A is $m > n$.

Least Norm problems

- Basic **Least Norm Problems** have the form minimize $||x||$ subject to $Ax = b$ where $A \in R^{m \times n}$ and $b \in \boldsymbol{R}^m$ and $x \in \boldsymbol{R}^n$ is the variable.
- They have the following properties:
  - The problem only becomes interesting when $m < n$.
  - With x_0 as any solution to $Ax = b$ and $Z \in \boldsymbol{R}^{n \times k}$ as any matrix whose columns are a basis for the *nullspace* of A. It can be rewritten as $||x_0 + Zu||$ with the variable $u \in \boldsymbol{R}^k$.

Interpretation and Example

- Estimation: $b = Ax$ are given measurements of x, x' is the smallest estimate consistent with measurements.

Examples of the Least Norm Problems

Least squares solution of linear equations: Minimize $\| x \|_2^2$ subject to $Ax = b$. This can be solved analytically by $2x^* + A^T v^* = 0, Ax^* = b$ which is a pair of linear equations and can be readily solved. From the first equation we can see that $x^* = -(1/2)A^T v^*$, then substituting this into the second equation we get $-(1/2)AA^T v^* = b$, and conclude that $v^* = -2(AA^T)^{-}1b$ and $x^* = A^T (AA^T)^{-}1b$.

Bi-criterion formulation in Regularization

- **Bi-criterion Problem** is a vector optimization problem with two objectives:

$$\text{minimize}\ (w.r.t.\ R_+^2)\ (\|Ax - b\|, \|x\|)$$

- The *optimal trade-off curve* of $\|Ax - b\|$, the residual, versus $\|x\|$ shows how large one of the objectives must be to have the other one small.
- Solution: $x = A^{-1}b$.

Regularization

- The goal of **Regularization** is to find a good approximation of $Ax \approx b$ with a small vector x.
- Adds an *extra term* or *parameter associated* with the norm of x.
- Examples:
  - minimize $||Ax - b|| + \gamma||x||$, where $\gamma > 0$.
  - minimize $||Ax - b||^2 + \delta||x||^2$ where $\delta > 0$.

Tikhonov regularization

- minimize

$$\|Ax - b\|_2^2 + \delta\|x\|_2^2$$

where $\delta > 0$.

- Can be solved as a *least-squares problem*:

$$\text{minimize} \left\| \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2$$

- Solution: $x = (A^T A + \delta \mathrm{I})^{-1} A^T b$

Other Regularization Methods

- **Smoothing regularization**
  - Replaces $\|x\|$ with $\|Dx\|$
  - The matrix D represents the *approximate second-order differentiation operator term* and measures the smoothness of x.
- **ℓ1-norm regularization**
  - minimize $\|Ax - b\|_2 + \gamma\|x\|_1$ where $\gamma > 0$.
  - Can approximate the *optimal trade-off curve* between $\|Ax - b\|_2$ and cardinality *card(x)* of the vector x.

Reconstruction and Smoothing

- **Reconstruction** is the process of forming an estimate $\hat{x}$ of an original signal represented by a vector $x \in R^n$
- Usually assumed that $x_i \approx x_{i+1}$.
- x is *corrupted* by an additive noise v:

$$x_{cor} = x + v$$

- **Smoothing** happens as an operation on x_{cor} to produce $\hat{x}$.
- The reconstruction problem:

$$\text{minimize } (w.r.t. R_+^2)\ (\|\hat{x} - x_{cor}\|_2, \varphi(\hat{x}))$$

- Measures the lack of *smoothness* of the estimate $\hat{x}$.

Quadratic smoothing

- The **Quadratic Smoothing Function**:

$$\varphi_{quad}(x) = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 = \|Dx\|_2^2$$

where $D \in R^{(n-1)\times n}$ is the *bidiagonal* matrix.

- The bi-criterion problem minimizes:

$$\|\hat{x} - x_{cor}\|_2^2 + \delta\|D\hat{x}\|_2^2$$

where δ > 0.

- Solution: $\hat{x} = (I + \delta \mathrm{D}^T D)^{-1} x_{cor}$

Total variation reconstruction

- Used when the original signal is very smooth, and the noise is rapidly varying.
- Based on the smoothing function:

$$\varphi_{tv}(\hat{x}) = \sum_{i=1}^{n-1} |\hat{x}_{i+1} - \hat{x}_i| = \|D\hat{x}\|_1$$

Robust Approximation

Recall the simplest norm approximation problem:

$$\text{minimize} \qquad ||Ax - b||$$

It is always desirable to take the possible *variation* of the matrix $A \in \mathbb{R}^{m \times n}$ into consideration, which is called the **Robust Approximation**.

- Stochastic Robust Approximation Problem (SRAP)
- Worst- Case Robust Approximation Problem(WCRAP)
- Properties of SRAP and WCRAP
- Examples

Stochastic Robust Approximation Problem

- **Idea**

Decompose the random matrix A in $\mathbb{R}^{m\times n}$ in the following way,

$$A = \bar{A} + U.$$

where $\bar{A} = \mathrm{E}(A)$ and U is a random matrix in $\mathbb{R}^{m\times n}$ with zero mean.

- **Objective**

$$\textbf{minimize} \quad \mathbf{E}||\boldsymbol{Ax} - \boldsymbol{b}||.$$

Worst- Case Robust Approximation Problem

- **Idea**

Develop a *nonempty* and *bounded* set $\mathcal{A}$, s.t. $A \in \mathcal{A} \subseteq \mathbb{R}^{m\times n}$, and take account the variation in *A* by considering the worst-case error of a candidate solution, x, given by

$$e_{wc}(x) = \sup\{||Ax - b|| \; |A \in \mathcal{A}\}$$

- **Objective**

$$\mathbf{minimize} \quad \boldsymbol{e_{wc}(x) = \sup\{||Ax - b|| \; |A \in \mathcal{A}\}}.$$

Properties of SRAP and WCRAP

- Both of them are always *convex problems*.
- Most of SRAP are *intractable* due to the difficulty in evaluating the objective and its derivative.
- The tractability of WCRAP depends on the *norm* used and the *uncertainty* set $\mathcal{A}$.

Example 1：Least-Square Problem of SRAP

- **Requirement**

  $\bar{A}$ and the covariance matrix of U^T are known and given by

$$P = \mathrm{E}[U^T U] = \mathrm{E}[A^T A]$$

- **Objective**

$$\text{minimize} \quad \mathrm{E}||Ax - b||_2^2$$

$$\rightarrow \qquad \text{minimize } ||\bar{\mathrm{A}}\mathrm{x} - \mathrm{b}||_2^2 + ||\sqrt{\mathrm{P}}x||_2^2$$

- **Comments**

  i. It's a *tractable* case of *SRAP*.

  ii. It has a solution as $\mathrm{x}^* = (\bar{A}^T\bar{A} + P)^{-1}\bar{A}^T b$, which corresponds to *Tikhonov regularization* with data matrix $\bar{\mathrm{A}}$ and $\Gamma = \sqrt{P}$.

Example 2：Sum-of-Norm Problem

- **Requirement**

  Data matrix A have only *finite values*, that is,

$$\Pr\left(A = A_i\right) = p_i\,, \quad i = 1, \cdots\cdots, k.$$

where $A_i \in \mathbb{R}^{m \times n}$, $1^T p = 1$, $p \succcurlyeq 0$.

- **Objective**

$$\text{minimize} \quad p_1 ||A_1 x - b|| + \cdots\cdots + p_k ||A_k x - b||.$$

- **Comments**

  i. It's a tractable *SRAP*.

  ii. Converted into a *LP* if l_∞-norm or l_1-norm are used.

Example 3: Norm Bound Error Problem

- **Requirement**

$\mathcal{A} = \{\bar{A} + U \mid ||U|| \le a\}$ is a norm ball, where $|| \bullet ||$ is a norm on $\mathbb{R}^{m\times n}$.

- **Objective** (if the Euclidean norm on $\mathbb{R}^m$ and maximum singular value norm on $\mathbb{R}^{m\times n}$ are used)

$$\text{minimize} \quad ||\bar{A}x - b||_2 + a||x||_2$$

$$\rightarrow \quad \text{minimize} \quad ||\bar{A}x - b||_2^2 + \delta||x||_2^2$$

- **Comments**

  i. If particular norms are used, the norm bound error problem is converted into a *regularized least square problem*.

  ii. Vice versa, we have another aspect to consider the regularized least square problem.

Example 4: Finite Set of WCRAP

- **Requirement**

  $\mathcal{A}$ is a finite set, $\mathcal{A} = \{A_1, \cdots\cdots, A_k\}$ or $\mathcal{A} = \mathrm{conv}(\{A_1, \cdots\cdots, A_k\})$.

- **Objective**

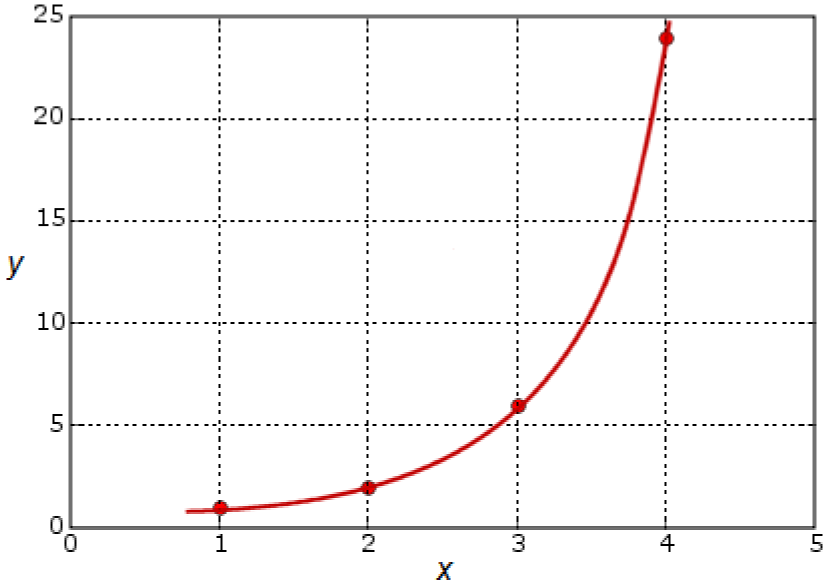
$$\text{minimize} \qquad \max ||A_i x - b||.$$

- **Comment**

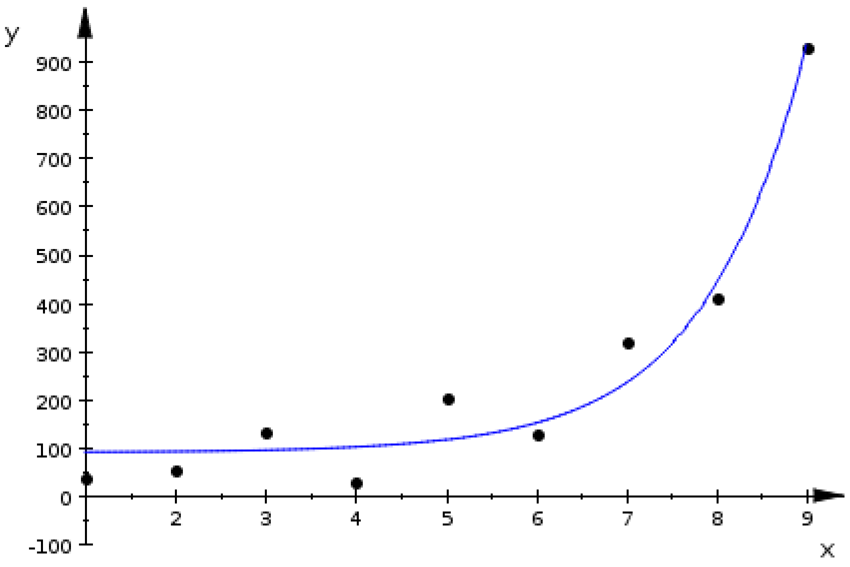
  Converted into a *LP* if l_∞-norm or l_1-norm are used.

Interpolation or the fit?

Interpolation



Fit



Problem statement

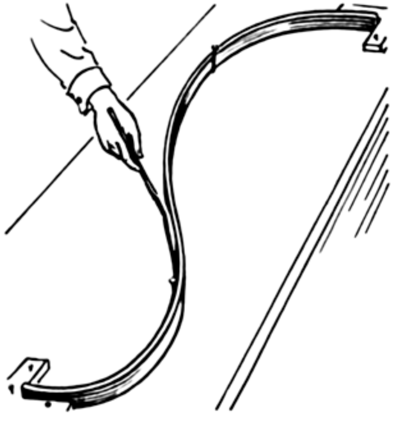
- Data points

$$(u_1, y_1), \ldots, (u_m, y_m)$$

- Approximating function

$$f(u) = x_1 f_1(u) + \cdots + x_n f_n(u)$$

- The family $\{f_1, \ldots, f_n\}$ (polynomials, trigonometric functions, splines, etc.)



Constraints

- Interpolation conditions

$$f(u_i) = y_i, \qquad i = 1, \dots, m,$$

- Lipschitz constraint

$$|f(u_j) - f(u_k)| \leq L \|u_j - u_k\|$$

- Nonnegativity
- Derivative / Integral

Types of problems

- Least-square norm $$\text{minimize} \sum_{i=1}^{m} (f(u_i) - y_i)^2$$

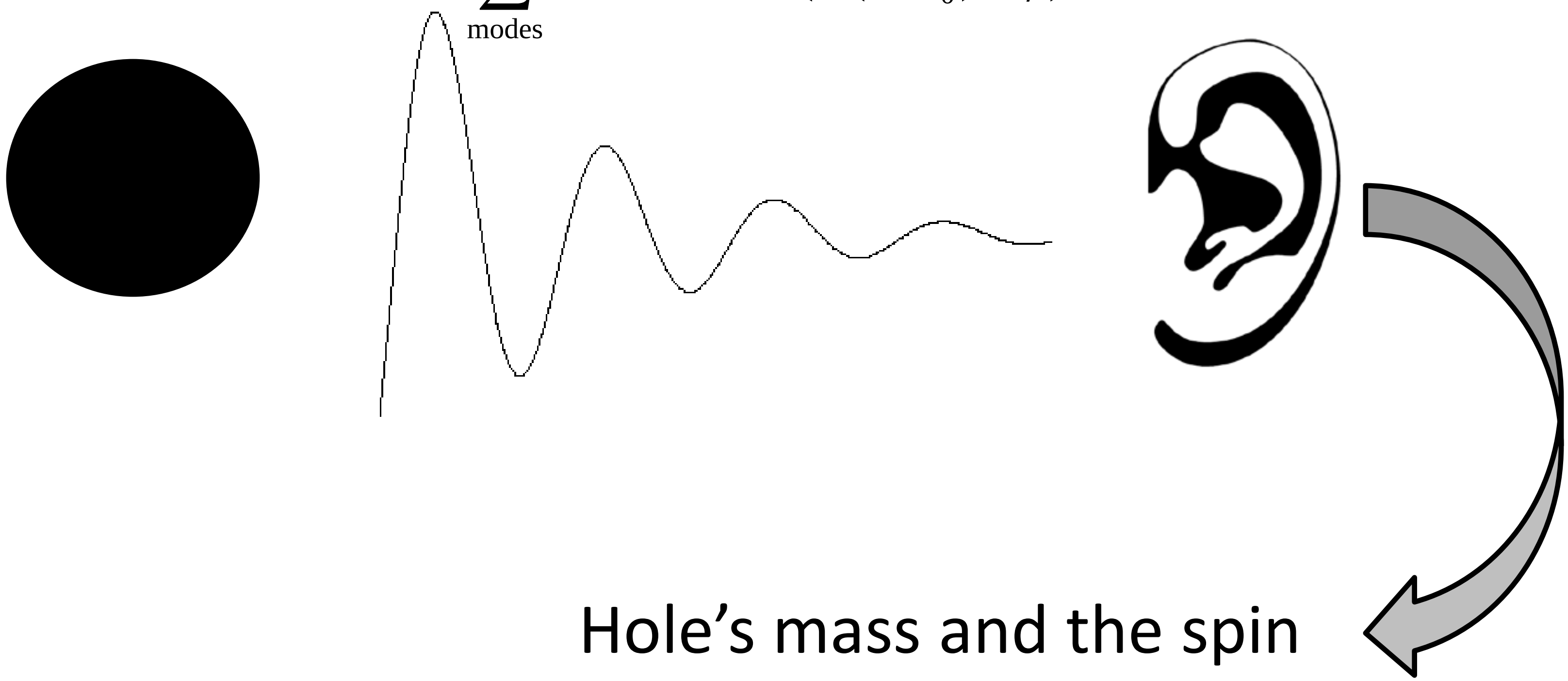
- Basis pursuit $$\text{minimize} \sum_{i=1}^{m} (f(u_i) - y_i)^2 + \gamma \|x\|_1,$$

- Fitting with a convex function

$$\text{minimize} \sum_{i=1}^{m} (f(u_i) - y_i)^2$$
$$\text{subject to } f: \mathbf{R}^{\mathbf{k}} \to \mathbf{R} \text{ is convex}$$

Extracting info from black hole radiation

$$\sum_{\text{modes}} A e^{-(t-t_0)/\tau} \cos(\omega(t-t_0)+\phi)$$



Hole’s mass and the spin

Bibliography

1. http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf, chapter 6, Apri13,2013.
2. *Robust Approximation*. Nishakova Irina, Seminar Optimierung
3. *Applications of Convex Optimization in Signal Processing and Communications*: Lecture 14. Andre Tkacenko, Signal Processing Research Group Jet Propulsion Laboratory. May 17, 2012.
4. http://en.wikipedia.org/wiki/Flat_spline, April 13,2013
5. http://www.systems.caltech.edu/dsp/ee150_acospc/lectures/EE_150_Lecture14 April 2013