ELEG 867 - Compressive Sensing and Sparse Signal Representations

Introduction to Matrix Completion and Robust PCA

Gonzalo Garateguy
Depart. of Electrical and Computer Engineering
University of Delaware

Fall 2011

Matrix Completion Problems - Motivation

Recomender Systems

NETFLIX

amazon.com

iTunes

Google

BARNES&NOBLE BOOKSELLERS

facebook

Items

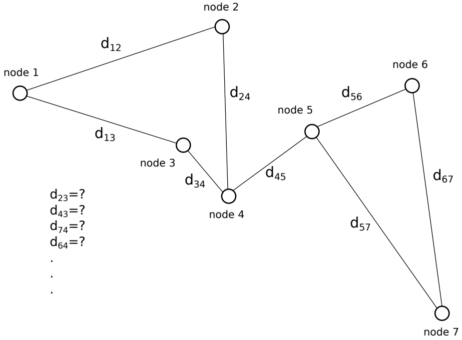
User 1	x	x	?	?	x	x
User 2	?	?	x	x	?	?
.	?	x	?	x	x	?
.	x	?	?	x	?	x
.	x	?	x	?	?	x
.	?	x	?	?	x	?
.	?	?	x	x	x	?
User n	x	x	?	?	?	x

- Collaborative filtering (Amazon, last.fm)
- Content based (Pandora, www.nanocrowd.com)
- Netflix prize competition boosted interest in the area

http://www.ima.umn.edu/videos/index.php?id=1598
http://sahd.pratt.duke.edu/Videos/keynote.html

Matrix Completion Problems - Motivation

Sensor location estimation in Wireless Sensor Networks



Distance matrix

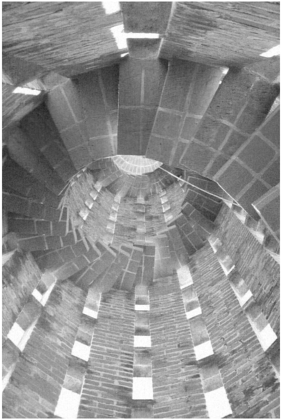
	1	2	3	4	5	6	7
1	0	$d_{1,2}$	$d_{1,3}$	?	?	?	?
2	$d_{2,1}$	0	?	$d_{2,4}$	?	?	?
3	$d_{3,1}$	?	0	$d_{3,4}$	?	?	?
4	?	$d_{4,2}$	$d_{4,3}$	0	$d_{4,5}$	?	?
5	?	?	?	$d_{5,4}$	0	$d_{5,6}$	$d_{5,7}$
6	?	?	?	?	$d_{6,5}$	0	$d_{6,7}$
7	?	?	?	?	$d_{7,5}$	$d_{7,6}$	0

- The problem is to find the positions of the sensors in R^2 given the partial information about relative distances
- A distance matrix like this has rank 2 in R^2
- For certain types of graphs the problem can be solved if we know the whole distance matrix

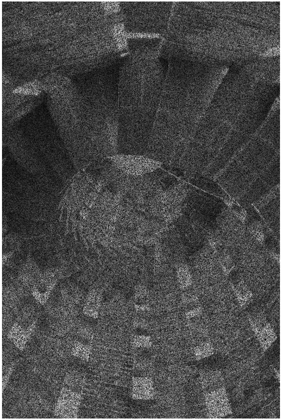
Matrix Completion Problems - Motivation

Image reconstruction from incomplete data

Reconstructed image



Incomplete image 50% of the pixels



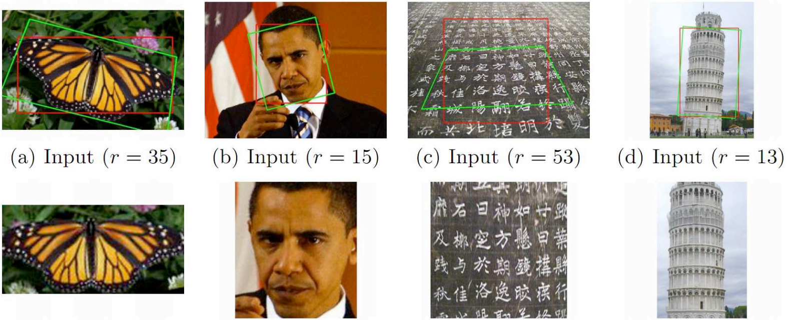
Robust PCA - Motivation

Foreground identification for surveillance applications

E.J. Candes, X. Li, Y. Ma, and Wright, J. "Robust principal component analysis?" `http://arxiv.org/abs/0912.3599`

Robust PCA - Motivation

Image alignment and texture recognition



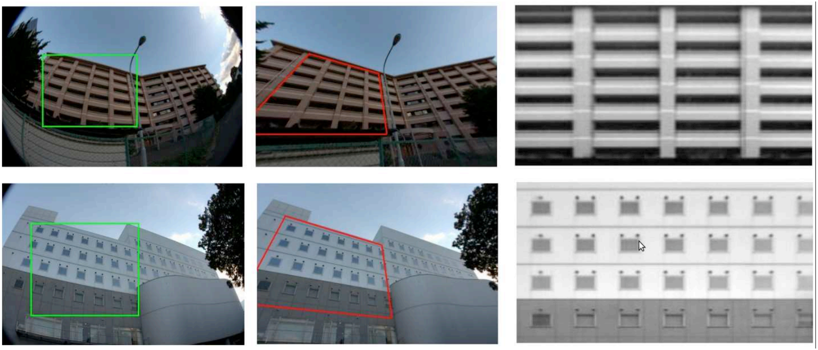
(a) Input ($r = 35$) (b) Input ($r = 15$) (c) Input ($r = 53$) (d) Input ($r = 13$)

(e) Output ($r = 14$) (f) Output ($r = 8$) (g) Output ($r = 19$) (h) Output ($r = 6$)

Z. Zhang, X. Liang, A. Ganesh, and Y. Ma, “TILT: transform invariant low-rank textures” Computer Vision–ACCV 2010

Robust PCA - Motivation

Camera calibration with radial distortion



J. Wright, Z. Lin, and Y. Ma "Low-Rank Matrix Recovery: From Theory to Imaging Applications" Tutorial presented at International Conference on Image and Graphics (ICIG), August 2011

Motivation

Many other applications

- System Identification in control theory
- Covariance matrix estimation
- Machine Learning
- Computer Vision

Videos to watch

Matrix Completion via Convex Optimization: Theory and Algorithms by Emmanuel Candes
`http://videolectures.net/mlss09us_candes_mccota/`

Low Dimensional Structures in Images or Data by Yi Ma, Workshop in Signal Processing with Adaptive Sparse Structured Representations (June 2011)

`http://ecos.maths.ed.ac.uk/SPARS11/YiMa.wmv`

Problem Formulation

Matrix completion

$$\begin{aligned} \text{minimize} \quad & \text{rank}(\mathbf{A}) \qquad (1) \\ \text{subject to} \quad & A_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{aligned}$$

Robust PCA

$$\begin{aligned} \text{minimize} \quad & \text{rank}(\mathbf{A}) + \lambda ||E||_0 \qquad (2) \\ \text{subject to} \quad & A_{ij} + E_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{aligned}$$

- Very hard to solve in general without any asumptions, some times NP hard.
- Even if we can solve them, are the solutions always what we expect?
- Under wich conditions we can have exact recovery of the real matrices?

Outline

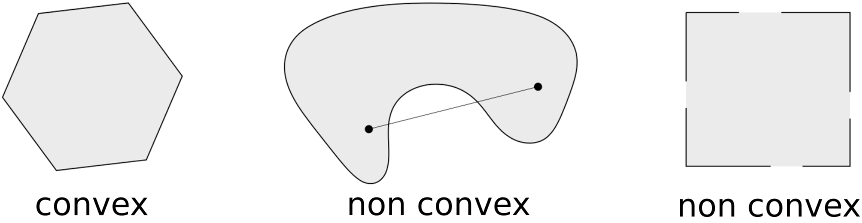
- Convex Optimization concepts
- Matrix Completion
  - Exact Recovery from incomplete data by convex relaxation
  - ALM method for Nuclear Norm Minimization
- Robust PCA
  - Exact Recovery from incomplete data and corrupted data by convex relaxation
  - ALM method for Low rank and Sparse separation

Convex sets and Convex functions

Convex set

A set C is convex if the line segment between any two points in C lies in C. For any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$ we have

$$\theta x_1 + (1 - \theta) x_2 \in C.$$



Convex sets and Convex functions

Convex combination

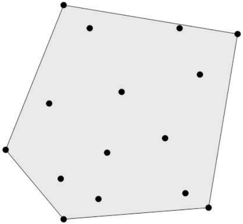
A convex combination of k points $x_1, .., x_k$ is defined as

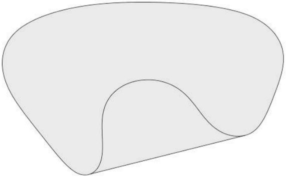
$$\theta_1 x_1 + ... + \theta_k x_k \ , \ \text{where} \ \theta_i \geq 0 \ \text{and} \ \theta_1 + ... + \theta_k = 1$$

Convex hull

The convex hull of C is the set of all convex conbinations of points in C

$$\textbf{conv}\ C = \{\theta_1 x_1 + ... + \theta_k x_k | x_i \in C, \theta_i \geq 0, i = 1, ..., k, \theta_1 + ... + \theta_k = 1\}$$





Convex sets and Convex functions

Operations that preserve convexity

Intersection

If $S1$ and $S2$ are convex, then $S_1 \bigcap S_2$ is convex.
In general if S_α is convex for every $\alpha \in \mathcal{A}$, then $\bigcap_{\alpha \in \mathcal{A}} S_\alpha$ is convex.

Subspaces, affine sets and convex cones are therefore closed under arbitrary intersections.

Affine functions

Let $f : R^n \to R^m$ be affine, $f(x) = Ax + b$, where $A \in R^{m \times n}$ and $b \in R^m$. If $S \subseteq R^n$ is convex, then the image of S under f

$$f(S) = \{f(x) | x \in S\}$$

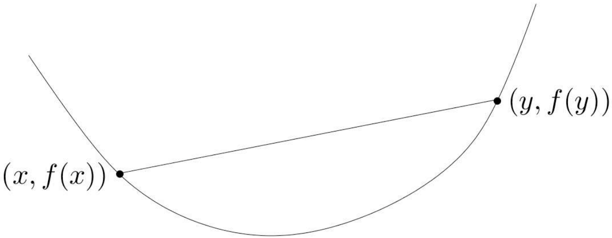
is convex

Convex functions

A function $f : R^n \rightarrow R$ is convex if **dom**f is a convex set and if for all $x, y \in$ **dom**f, and θ with $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta f(y))$$

we say that f is strictly convex if the strict intequality holds whenever $x \neq y$ and $0 < \theta < 1$



Operations that preserve convexity

Composition with an affine mapping

Suppose $f : R^n \to R, A \in R^{n\times m}$ and $b \in R^n$. Define $g : R^m \to R$ by

$$g(x) = f(Ax + b)$$

with $\mathbf{dom} g = \{x | Ax + b \in \mathbf{dom} f\}$. Then if f is convex, so is g.

Pointwise maximum

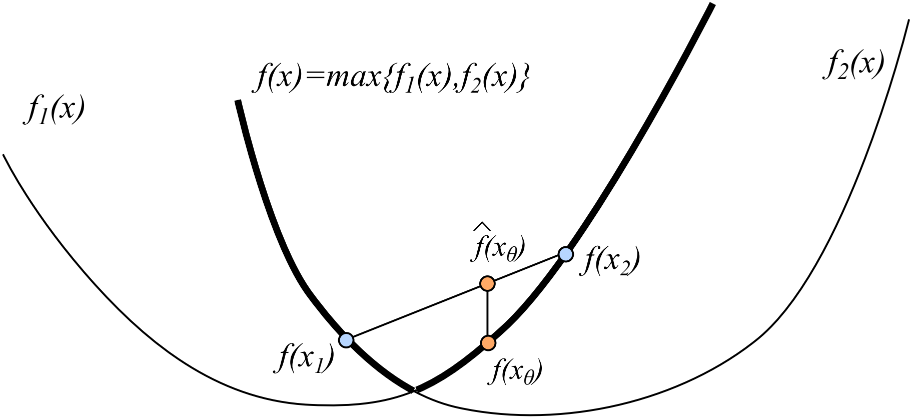
if f_1 and f_2 are convex functions then their pointwise maximum f defined by

$$f(x) = \max\{f_1(x), f_2(x)\}$$

with $\mathbf{dom} f = \mathbf{dom} f_1 \cap \mathbf{dom} f_2$ is also convex. This also extend to the case where $f_1, ..., f_m$ are convex, then

$$f(x) = \max\{f_1(x), ..., f_m(x)\}, \quad \text{is also convex}$$

Pointwise maximum of convex functions



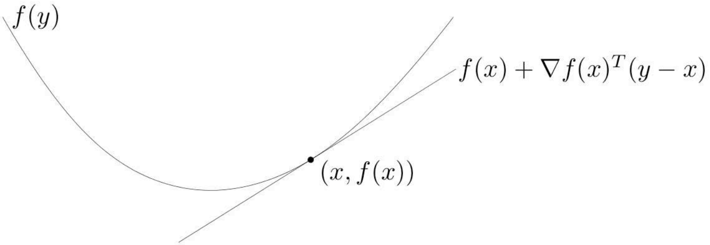
Convex sets and Convex functions

Convex differentiable functions

If f is differentiable (*i.e.* its gradient ∇f exist at each point in $\mathbf{dom} f$). Then f is convex if and only if $\mathbf{dom} f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

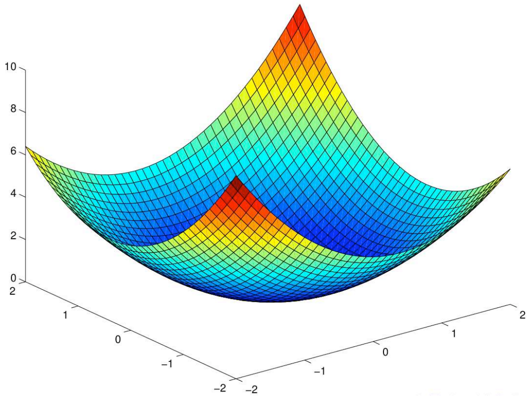
holds for all $x, y \in \mathbf{dom} f$.



Second order conditions

If f is twice differentiable, *i.e.* its Hessian $\nabla^2 f$ exist at each point in **dom**f. Then f is convex if and only if **dom**f is convex and its Hessian is positive semidefinite for all $x \in$**dom**f

$$\nabla^2 f(x) \succeq 0$$



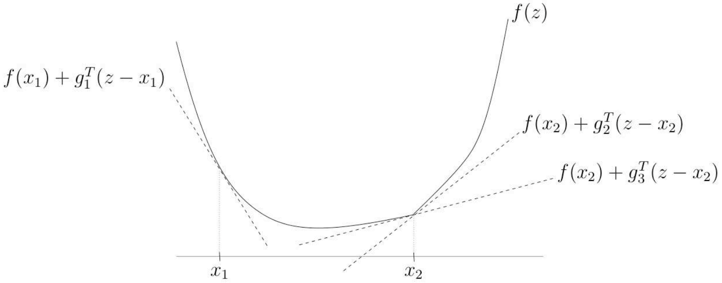
Convex non-differentiable functions

The concept of gradient can be extended to non-differentiable functions introducing the subgradient

Subgradient of a function

A vector $g \in R^n$ is a *subgradient* of $f : R^n \to R$ at $x \in \textbf{dom}f$ if for all $z \in \textbf{dom}f$

$$f(z) \geq f(x) + g^T(z - x)$$



Subgradients

Observations

- If f is convex and differentiable, then its gradient at x , $\nabla f(x)$ is its only subgradient

Subdifferentiable functions

A function f is called subdifferentiable at x if there exist at least one subgradient at x

Subdifferential at a point

The set of subgradients of f at the point x is called the subdifferential of f at x, and is denoted $\partial f(x)$

Subdifferentiability of a function

A function f is called subdifferentiable if it is subdifferentiable at all $x \in \mathbf{dom} f$

Basic properties

Existence of the subgradient of a convex function

If f is convex and $x \in \textbf{int dom} f$, then $\partial f(x)$ is nonempty and bounded.

The subdifferential $\partial f(x)$ is always a closed convex set, even if f is not convex. This follows from the fact that it is the intersection of an infinite set of halfspaces

$$\partial f(x) = \bigcap_{z \in \textbf{dom} f} \{g | f(z) \geq f(x) + g^T(z - x)\}.$$

Basic properties

Nonnegative scaling

For $\alpha \geq 0$, $\partial(\alpha f)(x) = \alpha \partial f(x)$

Subgradient of the sum

Given $f = f_1 + ... + f_m$, where $f_1, ..., f_m$ are convex functions, the subgradient of f at x is given by $\partial f(x) = \partial f_1(x) + ... + \partial f_m(x)$

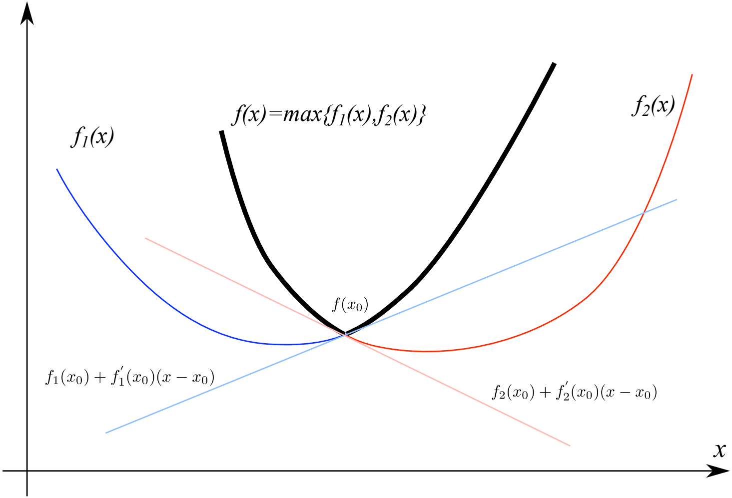
Affine transformations of domain

Suppose f is convex, and let $h(x) = f(Ax + b)$. Then $\partial h(x) = A^T \partial f(Ax + b)$.

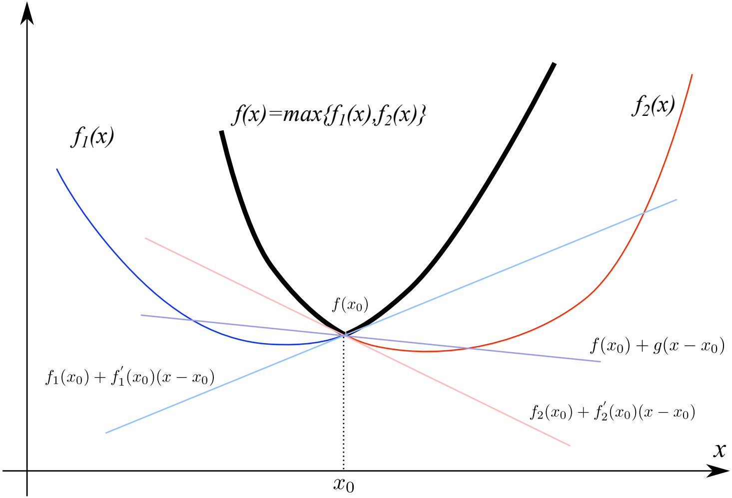
Pointwise maximum

Suppose f is the pointwise maximum of convex functions $f_1, ..., f_m$,
$f(x) = \max_{i=1,...,m} f_i(x)$, then $\partial f(x) = \mathbf{Co} \cup \{\partial f_i(x) | f_i(x) = f(x)\}$

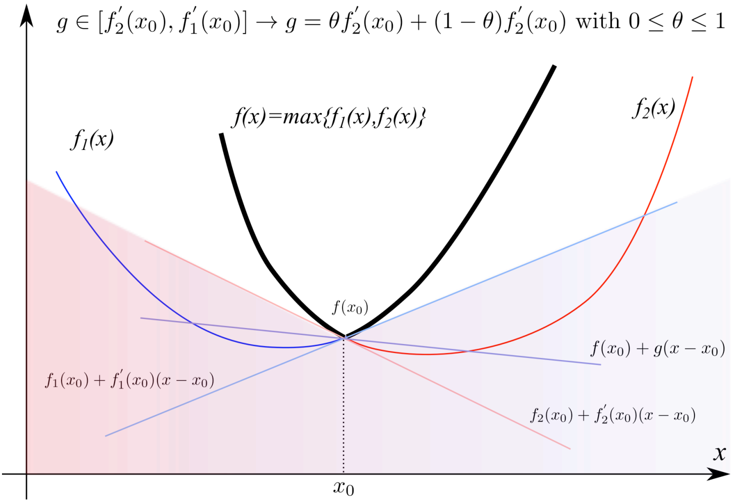
Subgradient of the pointwise maximum of two convex functions



Subgradient of the pointwise maximum of two convex functions



Subgradient of the pointwise maximum of two convex functions



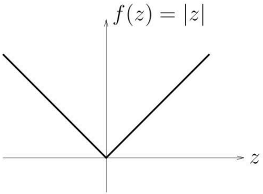
Examples

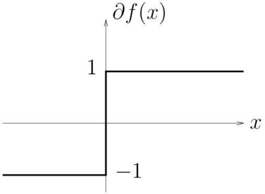
Conside the function $f(x) = |x|$. At x_0=0 , the subdiferential is defined by the inequality

$$\begin{aligned} f(z) \geq&\ f(x_0) + g(z - x_0), & \forall z \in \textbf{dom}\, f \\ |z| \geq&\ gz, & \forall z \in R \\ \partial f(0) =&\ \{g \mid g \in [-1\ ,\ 1]\} \end{aligned}$$

then for all x

$$\partial f(x) = \begin{cases} -1 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \\ \{g | g \in [-1, 1]\} & \text{for } x = 0 \end{cases}$$





Example: ℓ_1 norm

Consider $f(x) = \|x\|_1 = |x_1| + \cdots + |x_n|$, and note that f can be expressed as the maximum of 2^n linear functions

$$\|x\|_1 = \max\{\, f_1(x) \;, .., \; f_{2^n}(x) \,\}$$

$$\|x\|_1 = \max\{\, s_1^T x \;, .., \; s_{2^n}^T x \mid s_i \in \{-1, 1\}^n \,\}$$

The active functions $f_i(x)$ at x are the ones for wich $s_i^T x = \|x\|_1$. Then denoting

$$s_i = [s_{i,1}, ..., s_{i,n}]^T, \; s_{i,j} \in \{-1, 1\}$$

the set of indices of the active functions at x is

$$\mathcal{A}_x = \left\{ i \,\middle|\, \begin{array}{ll} s_{i,j} = -1 & \text{for } x_j < 0 \\ s_{i,j} = 1 & \text{for } x_j > 0 \\ s_{i,j} = -1 \text{ or } 1 & \text{for } x_j = 0 \end{array} \;, \text{for } j = 1, .., n \right\}$$

subgradient of the ℓ_1 norm

The subgradient of $\|x\|_1$ at a generic point x is defined by

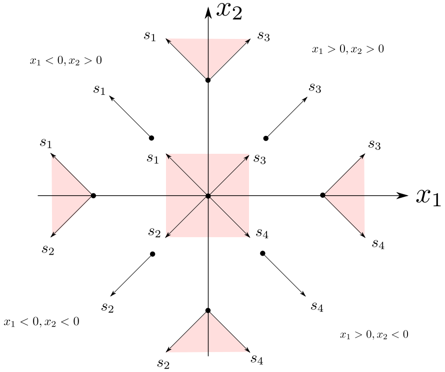
$$\begin{aligned}
\partial\|x\|_1 &= \mathbf{co} \cup \{ \partial f_i(x) \mid i \in \mathcal{A}_x \} \\
\partial\|x\|_1 &= \mathbf{co}\{ \nabla f_i(x) \mid i \in \mathcal{A}_x \} \\
\partial\|x\|_1 &= \mathbf{co}\{ s_i | i \in \mathcal{A}_x \} \\
\partial\|x\|_1 &= \{g | g = \textstyle\sum_{i\in\mathcal{A}_x} \theta_i s_i \ , \ \theta_i \geq 0 \ , \ \sum_i \theta_i = 1\}
\end{aligned}$$

or equivalently

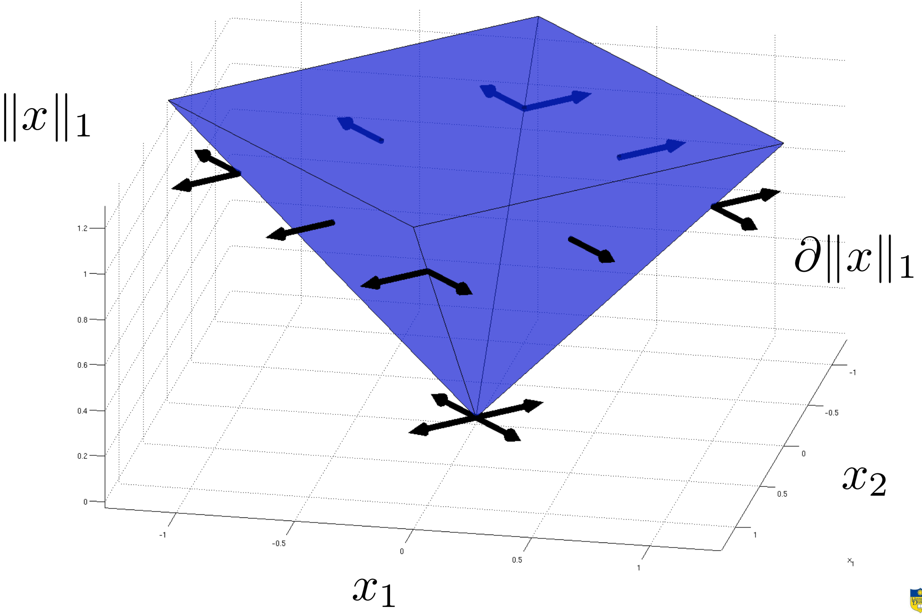
$$\partial\|x\|_1 = \left\{ g \;\middle|\; \begin{array}{lll} g_j = & -1 & \text{for } x_j < 0 \\ g_j = & 1 & \text{for } x_j > 0 \\ g_j = & \zeta \in [-1, 1] & \text{for } x_j = 0 \end{array} \right\}$$

ℓ_1 norm on R^2

in R^2 the set of subgradients are



$$
\begin{aligned}
s_1 &= [\ -1, \quad 1]^T \\
s_2 &= [\ -1, \ -1]^T \\
s_3 &= [\ \ 1, \quad 1]^T \\
s_4 &= [\ \ 1, \ -1]^T
\end{aligned}
$$



$\|x\|_1$
$\partial\|x\|_1$
x_1
x_2
1.2
1
0.8
0.6
0.4
0.2
0
-1
-0.5
0
0.5
1
-1
-0.5
0
0.5
1

Convex optimization problems

An optimization problem is convex if its objective is a convex function, the inequality constraints f_j are convex and the equality constraints h_j are affine

$$\begin{array}{rll} \underset{x}{\text{minimize}} & f_0(x) & \text{(Convex function)} \\ \text{s.t.} & f_i(x) \leq 0 & \text{(Convex sets)} \\ & h_j(x) = 0 & \text{(Affine)} \end{array}$$

or equivalently

$$\begin{array}{rll} \underset{x}{\text{minimize}} & f_0(x) & \text{(Convex function)} \\ \text{s.t.} & x \in C & C \text{ is a convex set} \\ & h_j(x) = 0 & \text{(Affine)} \end{array}$$

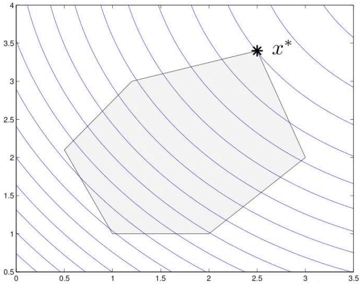
Theorem

If x^* is a local minimizer of a convex optimization problem, it is a global minimizer.

Optimality conditions

A point x^* is a minimizer of a convex function f if and only if f is subdifferentiable at x^* and

$$0 \in \partial f(x^*)$$



Convex optimization problems

Given the convex problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = \{1, ..., k\} \\ & h_j(x) = 0, \quad j = \{1, ..., l\} \end{aligned}$$

its Lagrangian function is defined as

$$\mathcal{L}(x, \lambda, \nu) = f_0(x) + \sum_{j=1}^{l} \lambda_j h_j(x) + \sum_{i=1}^{k} \nu_i f_i(x)$$

where $\nu_i \geq 0, \lambda_i \in R$

Augmented Lagrangian Method

Considering the problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{s.t.} \quad & x \in C \\ & \mathbf{h}(x) = 0 \end{aligned} \tag{3}$$

The augmented lagrangian is defined as

$$\mathcal{L}(x, \lambda, c) = f(x) + \lambda^T \mathbf{h}(x) + \frac{\mu}{2} \|\mathbf{h}(x)\|_2^2$$

where μ is a penalty parameter and λ is the multiplier vector

Augmented Lagrangian Method

The augmented lagrangian method consist of solving a sequence of problems of the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \mathcal{L}(x, \lambda_k, \mu_k) = f(x) + {\lambda_{\mathbf{k}}}^T \mathbf{h}(x) + \frac{\mu_k}{2} \|\mathbf{h}(x)\|_2^2 \\ \text{s.t.} \quad & x \in C \end{aligned}$$

where $\{\lambda_{\mathbf{k}}\}$ is a bounded sequence in R^l and $\{\mu_k\}$ is a penalty parameter sequence satisfying

$$0 < \mu_k < \mu_{k+1} \quad \forall\, k \,, \; \mu_k \rightarrow \infty$$

Augmented Lagrangian Method

The exact solution to problem (3) can be found using the following iterative algorithm

set $\rho > 1$

while not converged do

$$x_{k+1} = \underset{x \in C}{\operatorname{argmin}}\, \mathcal{L}(x, \lambda_{\mathbf{k}}, \mu_k)$$

solve (above)

$$\lambda_{\mathbf{k+1}} = \lambda_{\mathbf{k}} + \mu_k \mathbf{h}(x_{k+1})$$

$$\mu_k = \rho \mu_k$$

end while

Output x_k

Matrix completion

Optimization problem

$$\begin{array}{ll} \text{minimize} & \text{rank}(\mathbf{A}) \\ \text{subject to} & A_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{array} \tag{4}$$

- We look for the simplest explanation for the observed data
- Given enough number of samples, the likelihood of the solution to be unique should be high

Matrix completion

$$\begin{array}{ll} \text{minimize} & \text{rank}(A) \\ \text{subject to} & A_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{array}$$

- The minimization of the rank$(\cdot)$ function is a combinatorial problem, with exponential complexity in the size of the matrix!
- Need for a convex relaxation

$$\text{rank}(A) = ||diag(\Sigma)||_0 \qquad A = U\Sigma V^T$$
$$\Downarrow$$
$$||A||_* = ||diag(\Sigma)||_1$$

Convex relaxation

$$\begin{array}{ll} \text{minimize} & ||A||_* \\ \text{subject to} & A_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{array} \tag{5}$$

Matrix Completion

Nuclear Norm

The nuclear norm of a matrix $A \in R^{m \times n}$ is defined as $||A||_* = \sum_{i=1}^{r} \sigma_i(A)$, where $\{\sigma_i(A)\}_{i=1}^{r}$ are the elements of the diagonal matrix Σ from the SVD decomposition of $A = U \Sigma V^T$

Observations

- r = rank(A) can be $r < m, n$. If this is the case we say that the matrix is low rank
- the singular values $\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$ are obtained as the square root of the eigenvalues of $A^T A$ and are always $\sigma_i \geq 0$
- the left singular vectors U are the eigenvectors of AA^T
- the right singular vectors V are the eigenvectors of $A^T A$

Matrix Completion

Spectral Norm

The spectral norm of a matrix $A \in R^{m \times n}$ is defined as $\|A\|_2 = \sigma_{max}(A)$, where $\sigma_{max} = \max(\{\sigma_i(A)\}_{i=1}^{r})$

Dual Norm

Given an arbitrary norm $|| \cdot ||_\diamond$ in R^n, its dual norm $|| \cdot ||_\dagger$ is defined as

$$\|z\|_\dagger = sup\{z^T x \mid \|x\|_\diamond \leq 1\}$$

Observations

- The nuclear norm is the dual norm of the spectral norm

$$\|A\|_* = sup\{\mathbf{tr}(A^T X) | \|X\|_2 \leq 1\}$$

Matrix Completion

Convex relaxation of the rank

Convex envelope of a function

Let $f : \mathcal{C} \rightarrow R$ where $\mathcal{C} \subseteq R^n$. The convex envelope of f (on $\mathcal{C}$) is defined as the largest convex function g such that $g(x) \leq f(x)$ for all $x \in \mathcal{C}$

Theorem

The convex envelope of the function $\phi(X)$ =rank(X) on $\mathcal{C} = \{X \in R^{m\times n} | \|X\|_2 \leq 1\}$, is $\phi_{env}(X) = \|X\|_*$.

Observations

- The convex envelope of rank(X) on a the set $\{X | \|X\|_2 \leq M\}$ is given by $\frac{1}{M}\|X\|_*$
- By solving the heuristic problem we obtain a lower bound on the optimal value of the original problem (provided we can identify a bound M on the feasible set).

M. Fazel, H. Hindi and S. Boyd "A Rank Minimization Heuristic with Application to Minimum Order System Approximation" American Control Conference, 2001.

Matrix completion

Convex relaxation

$$\begin{aligned} \text{minimize} \quad & \|A\|_* \\ \text{subject to} \quad & A_{ij} = D_{ij} \quad \forall (i,j) \in \Omega \end{aligned} \tag{6}$$

- The original problem is now a problem with a non-smooth but convex function as the objective
- The remaining problem is the number of measurements and in which positions have to be taken in order to guarantee that the solution is equal to the matrix D?

Matrix completion

Which types of matrices can be completed exactly?
Consider the matrix

$$M = e_1.e_n^T = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

- Can it be recovered from 90 % of its samples ?
- Is the sampling set important?
- Which sampling sets work and which ones doesn't?

Matrix completion

Sampling set Ω

The sampling set Ω is defined as $\Omega = \{(i,j) \mid D_{ij} \text{ is observed } \}$

Consider

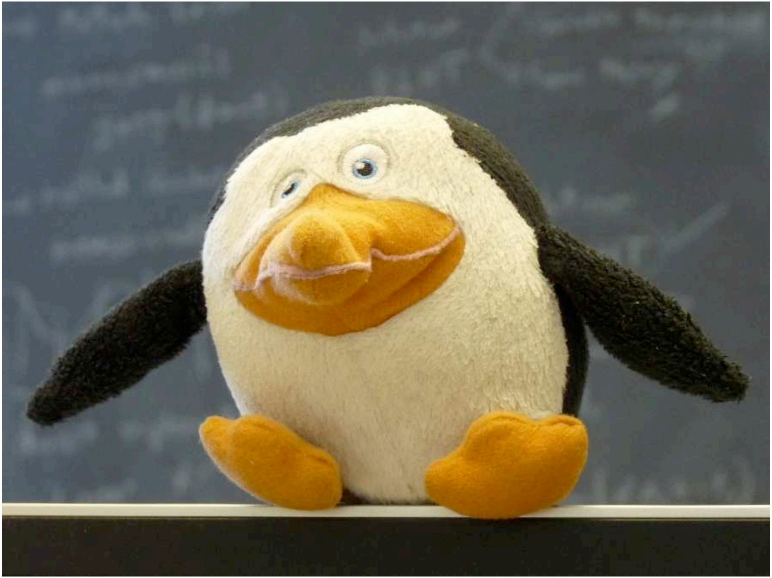
$$D = xy^T \quad x \in R^m, y \in R^n$$

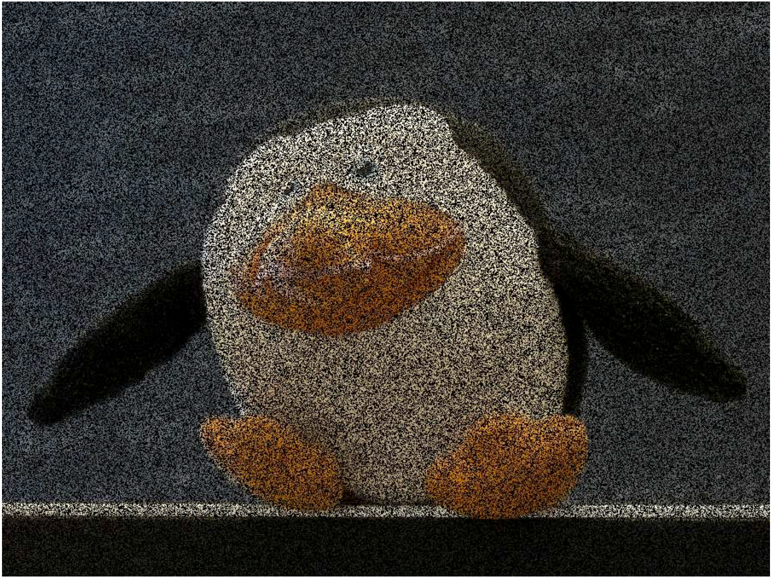
$$D_{ij} = x_i y_j$$

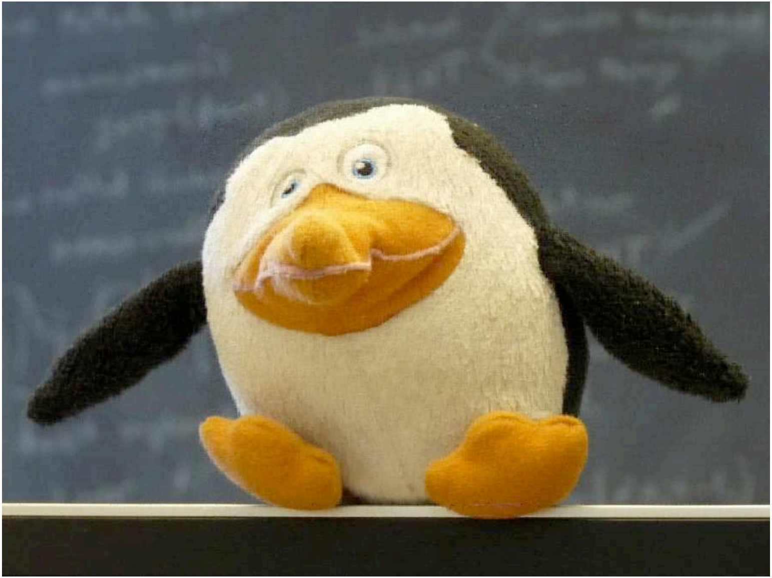
- If the sampling set avoids row i, then x_i can not be recovered by any method whatsoever

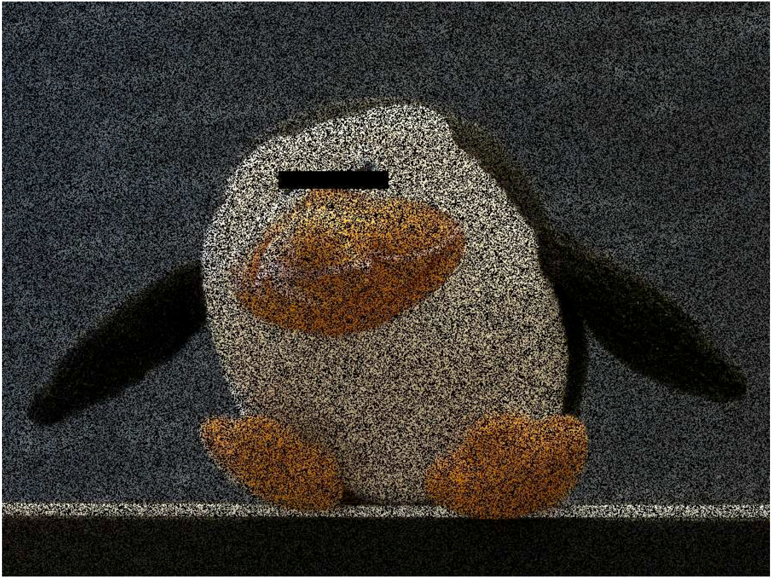
Observation

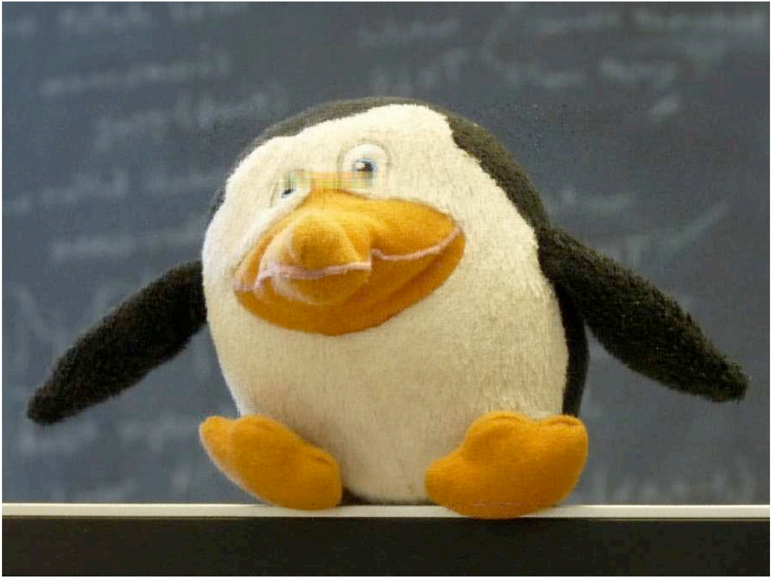
- No columns or rows from D can be avoided in the sampling set
- There is a need for a characterization of the sampling operator with respect to the set of matrices that we want to recover

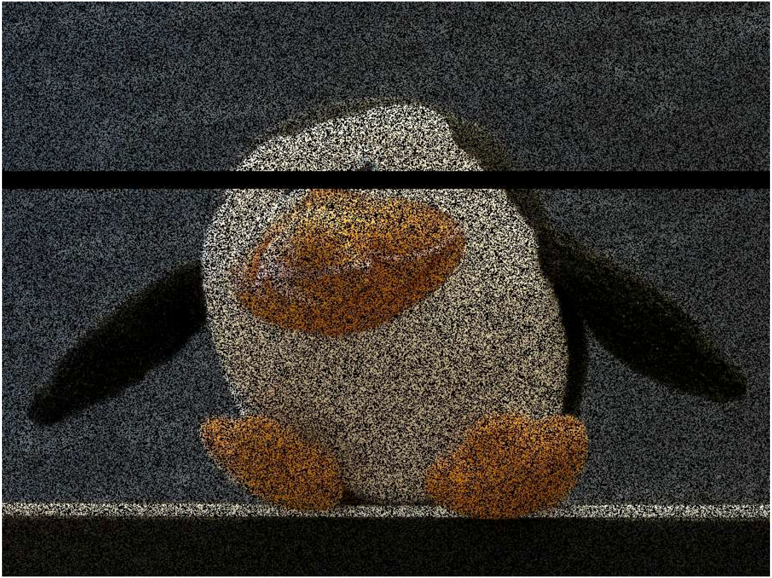


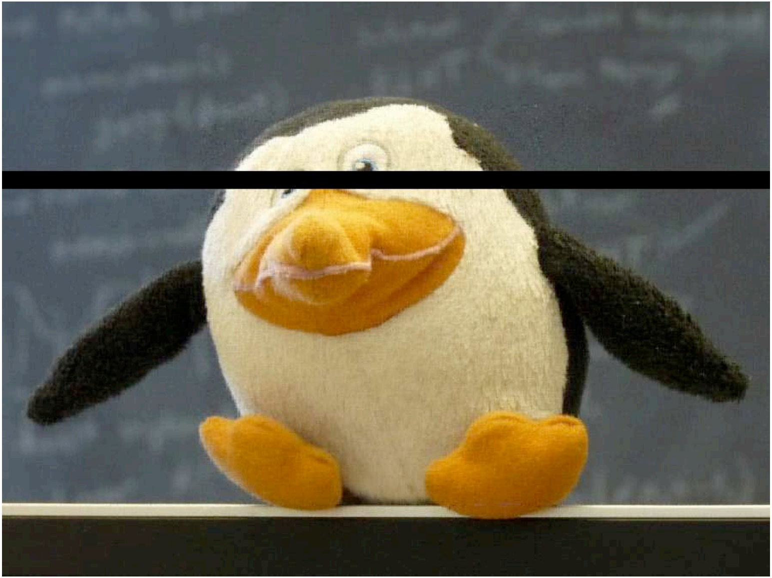






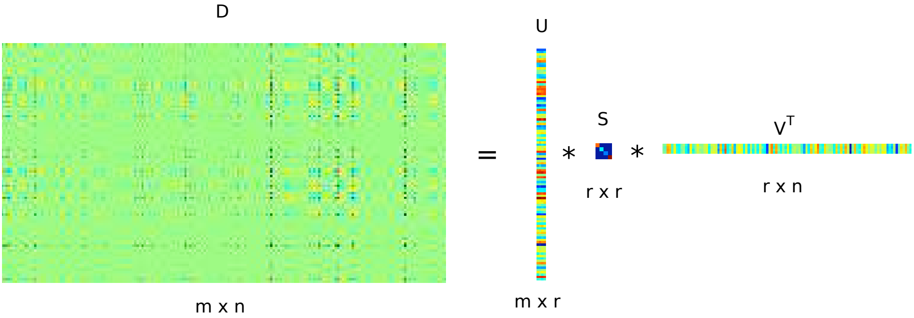






Matrix completion

- To recover a low rank matrix, this matrix cannot be in the null space of the sampling operator
- If the singular vectors of $D = USV^T$ are highly concentrated, then D is more likely to be in the null space of a given sampling operator



Matrix completion

Intuition

- the singular vectors need to be sufficiently spread, *i.e.* uncorrelated with the standar basis in order to minimize the number of observations needed to recover a low rank matrix

Coherence of a subspace

Let U be a subspace of R^n of dimension r and P_U be the orthogonal projection onto U. Then the coherence of U is defined to be

$$\mu(U) = \frac{n}{r} \max_{1 \leq i \leq n} \|P_U e_i\|^2$$

Observations

- The minimum value that $\mu(U)$ can achieve is 1 for example if U is spanned by vectors whos entries all have magnitude $1/\sqrt{n}$
- The largest possible value for $\mu(U)$ is n/r corresponding to a subspace that contains a standard basis element.

Matrix completion

μ_0 coherence

A matrix $D = \sum_{1 \leq k \leq r} \sigma_k u_k v_k^T$ is μ_0 coherent if for some positive μ_0 $\max(\mu(U), \mu(V)) \leq \mu_0$

μ_1 coherence

A matrix $D = \sum_{1 \leq k \leq r} \sigma_k u_k v_k^T$ has μ_1 coherence if

$$\|UV^T\|_\infty \leq \mu_1 \sqrt{r/mn}$$

for some $\mu_1 > 0$

Observation

- If D is μ_0 coherent then it is μ_1 coherent for $\mu_1 = \mu_0 \sqrt{r}$

Matrix completion

Theorem

Let $D \in R^{m\times n}$ of rank r be (μ_0, μ_1)-coherent and let $N = max(m,n)$. If we observe M entries of D with locations sampled uniformly at random. Then there exist constants C and c such that if

$$M \geq C \max(\mu_1^2, \mu_0^{1/2}\mu_1, \mu_0 N^{1/4}) Nr(\beta log N)$$

for some $\beta > 2$, then the minimizer of (6) is unique and equal to D with probability at least $1 - cn^{-\beta}$. If in addition $r \leq \mu_0^{-1} N^{1/5}$ then the number of observations can be improved to

$$M \geq C\mu_0 N^{6/5} r(\beta log N)$$

Candès, E.J. and Recht, B. "Exact matrix completion via convex optimization", Foundations of Computational Mathematics 2009

Matrix completion

Recovery performance

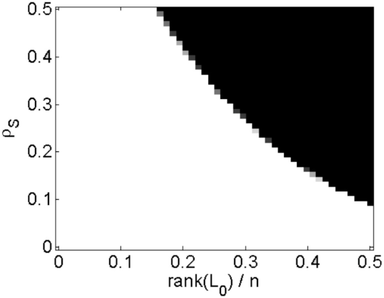


Figure: The x axis corresponds to $\text{rank}(A)/\min\{m, n\}$ and the y axis to $\rho_s = 1 - M/mn$ (probability that an entry is omited from the observations)

Emmanuel J. Candes, Xiaodong Li, Yi Ma, John Wright "Robust Principal Component Analysis?"

`http://arxiv.org/abs/0912.3599`

Matrix completion

Other bounds on number of meassurements and sampling operators

- Emmanuel J. Candes, Xiaodong Li, Yi Ma, John Wright "Rodbust Principal Component Analysis?"
  `http://arxiv.org/abs/0912.3599`
- Venkat Chandrasekaran, Sujay Sanghavi, Pablo A. Parrilo, Alan S. Willsky "Rank-Sparsity Incoherence for Matrix Decomposition"
  `http://arxiv.org/abs/0906.2220`
- Zihan Zhou, Xiaodong Li, John Wright, Emmanuel Candes, Yi Ma "Stable Principal Component Pursuit"
  `http://arxiv.org/abs/1001.2363`
- Raghunandan H. Keshavan, Andrea Montanari, Sewoong Oh "Matrix Completion from a Few Entries"
  `http://arxiv.org/abs/0901.3150`
- Sahand Negahban, Martin J. Wainwright "Restricted strong convexity and weighted matrix completion: Optimal bounds with noise"
  `http://arxiv.org/abs/1009.2118v2`
- Yonina C. Eldar, Deanna Needell, Yaniv Plan "Unicity conditions for low-rank matrix recovery"
  `http://arxiv.org/abs/1103.5479`