13. Douglas-Rachford method and ADMM

- Douglas-Rachford splitting method
- examples
- alternating direction method of multipliers
- image deblurring example
- convergence
Douglas-Rachford splitting algorithm

minimize \( f(x) + g(x) \)

\( f \) and \( g \) are closed convex functions

**Douglas-Rachford iteration:** start at any \( y^{(0)} \) and repeat for \( k = 1, 2, \ldots \),

\[
\begin{align*}
x^{(k)} &= \text{prox}_f(y^{(k-1)}) \\
y^{(k)} &= y^{(k-1)} + \text{prox}_g(2x^{(k)} - y^{(k-1)}) - x^{(k)}
\end{align*}
\]

- useful when \( f \) and \( g \) have inexpensive prox-operators
- \( x^{(k)} \) converges to a solution of \( 0 \in \partial f(x) + \partial g(x) \) (if a solution exists)
- not symmetric in \( f \) and \( g \)
Douglas-Rachford iteration as fixed-point iteration

• iteration on page 13-2 can be written as fixed-point iteration

\[ y^{(k)} = F(y^{(k-1)}) \]

where

\[ F(y) = y + \text{prox}_g(2\text{prox}_f(y) - y) - \text{prox}_f(y) \]

• \( y \) is a fixed point of \( F \) if and only if \( x = \text{prox}_f(y) \) satisfies \( 0 \in \partial f(x) + \partial g(x) \):

\[ y = F(y) \]

\[ \updownarrow \]

\[ 0 \in \partial f(\text{prox}_f(y)) + \partial g(\text{prox}_f(y)) \]

(proof on next page)
Proof.

\[ x = \text{prox}_f(y), \quad y = F(y) \]
\[ \quad \upepsilon \]
\[ x = \text{prox}_f(y), \quad x = \text{prox}_g(2x - y) \]
\[ \quad \upepsilon \]
\[ y - x \in \partial f(x), \quad x - y \in \partial g(x) \]

• therefore, if \( y = F(y) \), then \( x = \text{prox}_f(y) \) satisfies

\[ 0 = (y - x) + (x - y) \in \partial f(x) + \partial g(x) \]

• conversely, if \(-z \in \partial f(x)\) and \(z \in \partial g(x)\), then \( y = x - z \) is a fixed point of \( F \)
Equivalent form of DR algorithm

- start iteration on page 13-2 at $y$-update and renumber iterates

\[
y^{(k)} = y^{(k-1)} + \text{prox}_g(2x^{(k-1)} - y^{(k-1)}) - x^{(k-1)} \\
x^{(k)} = \text{prox}_f(y^{(k)})
\]

- switch $y$- and $x$-updates

\[
u^{(k)} = \text{prox}_g(2x^{(k-1)} - y^{(k-1)}) \\
x^{(k)} = \text{prox}_f(y^{(k-1)} + u^{(k)} - x^{(k-1)}) \\
y^{(k)} = y^{(k-1)} + u^{(k)} - x^{(k-1)}
\]

- make change of variables $w^{(k)} = x^{(k)} - y^{(k)}$

\[
u^{(k)} = \text{prox}_g(x^{(k-1)} + w^{(k-1)}) \\
x^{(k)} = \text{prox}_f(u^{(k)} - w^{(k-1)}) \\
w^{(k)} = w^{(k-1)} + x^{(k)} - u^{(k)}
\]
Scaling

algorithm applied to cost function scaled by $t > 0$

$$\text{minimize} \quad tf(x) + tg(x)$$

- algorithm of page 13-2

$$x^{(k)} = \text{prox}_{tf}(y^{(k-1)})$$
$$y^{(k)} = y^{(k-1)} + \text{prox}_{tg}(2x^{(k)} - y^{(k-1)}) - x^{(k)}$$

- algorithm of page 13-5

$$u^{(k)} = \text{prox}_{tg}(x^{(k-1)} + w^{(k-1)})$$
$$x^{(k)} = \text{prox}_{tf}(u^{(k)} - w^{(k-1)})$$
$$w^{(k)} = w^{(k-1)} + x^{(k)} - u^{(k)}$$

- the algorithm is not invariant with respect to scaling
- in theory, $t$ can be any positive constant; several heuristics exist for adapting $t$
Douglas-Rachford iteration with relaxation

- fixed-point iteration with relaxation

\[ y^{(k)} = y^{(k-1)} + \rho_k (F(y^{(k-1)}) - y^{(k-1)}) \]

\( 1 < \rho_k < 2 \) is overrelaxation, \( 0 < \rho_k < 1 \) underrelaxation

- algorithm of page 13-2 with relaxation

\[ x^{(k)} = \text{prox}_f(y^{(k-1)}) \]
\[ y^{(k)} = y^{(k-1)} + \rho_k (\text{prox}_g(2x^{(k)} - y^{(k-1)}) - x^{(k)}) \]

- algorithm of page 13-5

\[ u^+ = \text{prox}_g(x + w) \]
\[ x^+ = \text{prox}_f(x + \rho(u^+ - x) - w) \]
\[ w^+ = w + x^+ - x + \rho(x - u^+) \]
Primal-dual formulation

primal: minimize \( f(x) + g(x) \)
dual: maximize \(-g^*(z) - f^*(-z)\)

• use Moreau decomposition to simplify step 2 of DR iteration (page 13-2):

\[
x^{(k)} = \text{prox}_f(y^{(k-1)}) \\
y^{(k)} = x^{(k)} - \text{prox}_{g^*}(2x^{(k)} - y^{(k-1)})
\]

• make change of variables \( z^{(k)} = x^{(k)} - y^{(k)} \):

\[
x^{(k)} = \text{prox}_f(x^{(k-1)} - z^{(k-1)}) \\
z^{(k)} = \text{prox}_{g^*}(z^{(k-1)} + 2x^{(k)} - x^{(k-1)})
\]
Outline

- Douglas-Rachford splitting method
- examples
- alternating direction method of multipliers
- image deblurring example
- convergence
Sparse inverse covariance selection

\[
\text{minimize} \quad \operatorname{tr}(CX) - \log \det X + \gamma \sum_{i > j} |X_{ij}|
\]

variable is $X \in S^n$; parameters $C \in S_{++}^n$ and $\gamma > 0$ are given

Douglas-Rachford splitting

\[
f(X) = \operatorname{tr}(CX) - \log \det X, \quad g(X) = \gamma \sum_{i > j} |X_{ij}|
\]

\begin{itemize}
  \item $X = \text{prox}_{tf}(\hat{X})$ is positive solution of $C - X^{-1} + (1/t)(X - \hat{X}) = 0$
    easily solved via eigenvalue decomposition of $\hat{X} - tC$ (see homework)
  \item $X = \text{prox}_{tg}(\hat{X})$ is soft-thresholding
\end{itemize}
Spingarn’s method of partial inverses

Equality constrained convex problem \((f\) closed and convex; \(V\) a subspace)

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in V
\end{align*}
\]

Spingarn’s method: Douglas-Rachford splitting with \(g = \delta_V\) (indicator of \(V\))

\[
\begin{align*}
x^{(k)} &= \text{prox}_t f(y^{(k-1)}) \\
y^{(k)} &= y^{(k-1)} + P_V(2x^{(k)} - y^{(k-1)}) - x^{(k)}
\end{align*}
\]

Primal-dual form (algorithm of page 13-8):

\[
\begin{align*}
x^{(k)} &= \text{prox}_t f(x^{(k-1)} - z^{(k-1)}) \\
z^{(k)} &= P_{V^\perp}(z^{(k-1)} + 2x^{(k)} - x^{(k-1)})
\end{align*}
\]
Application to composite optimization problem

minimize \( f_1(x) + f_2(Ax) \)

\( f_1 \) and \( f_2 \) have simple prox-operators

- problem is equivalent to minimizing \( f(x_1, x_2) \) over subspace \( V \) where

\[
f(x_1, x_2) = f_1(x_1) + f_2(x_2), \quad V = \{(x_1, x_2) \mid x_2 = Ax_1\}
\]

- \( \text{prox}_{t f} \) is separable:

\[
\text{prox}_{t f}(x_1, x_2) = (\text{prox}_{t f_1}(x_1), \text{prox}_{t f_2}(x_2))
\]

- projection of \( (x_1, x_2) \) on \( V \) reduces to linear equation:

\[
P_V(x_1, x_2) = \begin{bmatrix} I & A \\ \end{bmatrix} (I + A^T A)^{-1}(x_1 + A^T x_2)
\]

\[
= \begin{bmatrix} x_1 \\ x_2 \\ \end{bmatrix} + \begin{bmatrix} A^T \\ -I \\ \end{bmatrix} (I + AA^T)^{-1}(x_2 - Ax_1)
\]
Decomposition of separable problems

\[
\begin{align*}
\text{minimize} & \quad \sum_{j=1}^{n} f_j(x_j) + \sum_{i=1}^{m} g_i(A_{i1}x_1 + \cdots + A_{in}x_n) \\
\end{align*}
\]

- same problem as page 12-17, but without strong convexity assumption
- we assume the functions \(f_j\) and \(g_i\) have inexpensive prox-operators

Equivalent formulation

\[
\begin{align*}
\text{minimize} & \quad \sum_{j=1}^{n} f_j(x_j) + \sum_{i=1}^{m} g_i(y_{i1} + \cdots + y_{in}) \\
\text{subject to} & \quad y_{ij} = A_{ij}x_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n \\
\end{align*}
\]

- prox-operator of first term requires evaluations of \(\text{prox}_{t f_j}\) for \(j = 1, \ldots, n\)
- prox-operator of 2nd term requires \(\text{prox}_{n t g_i}\) for \(i = 1, \ldots, m\) (see page 8-8)
- projection on constraint set reduces to \(n\) independent linear equations
Decomposition of separable problems

Second equivalent formulation: introduce extra splitting variables $x_{ij}$

\[
\begin{align*}
&\text{minimize} & & \sum_{j=1}^{n} f_j(x_j) + \sum_{i=1}^{m} g_i(y_{i1} + \cdots + y_{in}) \\
&\text{subject to} & & x_{ij} = x_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n \\
& & & y_{ij} = A_{ij}x_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n
\end{align*}
\]

- make first set of constraints part of domain of $f_j$:

\[
\tilde{f}_j(x_j, x_{1j}, \ldots, x_{mj}) = \begin{cases} 
  f_j(x_j) & x_{ij} = x_j, \quad i = 1, \ldots, m \\
  +\infty & \text{otherwise}
\end{cases}
\]

prox-operator of $\tilde{f}_j$ reduces to prox-operator of $f_j$

- projection on other constraints involves $mn$ independent linear equations
Outline

- Douglas-Rachford splitting method
- examples
- alternating direction method of multipliers
- image deblurring example
- convergence
Dual application of Douglas-Rachford method

Separable convex problem

\[
\begin{align*}
\text{minimize} & \quad f_1(x_1) + f_2(x_2) \\
\text{subject to} & \quad A_1 x_1 + A_2 x_2 = b
\end{align*}
\]

Dual problem

\[
\begin{align*}
\text{maximize} & \quad -b^T z - f_1^*(-A_1^T z) - f_2^*(-A_2^T z)
\end{align*}
\]

we apply the Douglas-Rachford method (page 13-5) to minimize

\[
\underbrace{b^T z + f_1^*(-A_1^T z)}_{g(z)} + \underbrace{f_2^*(-A_2^T z)}_{f(z)}
\]
Douglas Rachford applied to the dual

\[ u^+ = \text{prox}_{tg}(z + w), \quad z^+ = \text{prox}_{tf}(u^+ - w), \quad w^+ = w + z^+ - u^+ \]

**First line**: use result on page 10-7 to compute \( u^+ = \text{prox}_{tg}(z + w) \)

\[
\hat{x}_1 = \arg\min_{x_1} (f_1(x_1) + z^T(A_1x_1 - b) + \frac{t}{2}\|A_1x_1 - b + w/t\|_2^2)
\]

\[
u^+ = z + w + t(A_1\hat{x}_1 - b)
\]

**Second line**: similarly, compute \( z^+ = \text{prox}_{tf}(z + t(A_1\hat{x}_1 - b)) \)

\[
\hat{x}_2 = \arg\min_{x_2} (f_2(x_2) + z^T A_2x_2 + \frac{t}{2}\|A_1\hat{x}_1 + A_2x_2 - b\|_2^2)
\]

\[
z^+ = z + t(A_1\hat{x}_1 + A_2\hat{x}_2 - b)
\]

**Third line** reduces to \( w^+ = tA_2\hat{x}_2 \)
Alternating direction method of multipliers (ADMM)

1. minimize augmented Lagrangian over $x_1$

$$x_1^{(k)} = \arg\min_{x_1} \left( f_1(x_1) + (z^{(k-1)})^T A_1 x_1 + \frac{t}{2} ||A_1 x_1 + A_2 x_2^{(k-1)} - b||_2^2 \right)$$

2. minimize augmented Lagrangian over $x_2$

$$x_2^{(k)} = \arg\min_{x_2} \left( f_2(x_2) + (z^{(k-1)})^T A_2 x_2 + \frac{t}{2} ||A_1 x_1^{(k)} + A_2 x_2 - b||_2^2 \right)$$

3. dual update

$$z^{(k)} = z^{(k-1)} + t(A_1 x_1^{(k)} + A_2 x_2^{(k)} - b)$$

this the alternating direction method of multipliers or split Bregman method
Comparison with other multiplier methods

**Alternating minimization method** (page 12-22) with $g(y) = \delta_{\{b\}}(y)$

- same dual update, same update for $x_2$
- $x_1$-update in alternating minimization method is simpler:

$$x_1^{(k)} = \arg\min_{x_1} (f_1(x_1) + (z^{(k-1)})^T A_1 x_1)$$

- ADMM does not require strong convexity of $f_1$
- in theory, parameter $t$ in ADMM can be any positive constant

**Augmented Lagrangian method** (page 12-23) with $g(y) = \delta_{\{b\}}(y)$

- same dual update
- AL method requires joint minimization of the augmented Lagrangian

$$f_1(x_1) + f_2(x_2) + (z^{(k-1)})^T (A_1 x_1 + A_2 x_2) + \frac{t}{2} \|A_1 x_1 + A_2 x_2 - b\|^2_2$$

Douglas-Rachford method and ADMM
Application to composite optimization (method 1)

minimize \( f_1(x) + f_2(Ax) \)

• apply ADMM to

minimize \( f_1(x_1) + f_2(x_2) \)
subject to \( Ax_1 = x_2 \)

• augmented Lagrangian is

\[
 f_1(x_1) + f_2(x_2) + \frac{t}{2} \| Ax_1 - x_2 + z/t \|_2^2
\]

• \( x_1 \)-update requires (possibly nontrivial) minimization of

\[
 f_1(x_1) + \frac{t}{2} \| Ax_1 - x_2 + z/t \|_2^2
\]

• \( x_2 \)-update is evaluation of \( \text{prox}_{t^{-1} f_2} \)
Application to composite optimization (method 2)

introduce an extra ‘splitting’ variable $x_3$

\[
\text{minimize} \quad f_1(x_3) + f_2(x_2)
\]

\[
\text{subject to} \quad \begin{bmatrix} A \\ I \end{bmatrix} x_1 = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}
\]

• alternate minimization of augmented Lagrangian over $x_1$ and $(x_2, x_3)$

\[
f_1(x_3) + f_2(x_2) + \frac{t}{2} \left( \|Ax_1 - x_2 + z_1/t\|_2^2 + \|x_1 - x_3 + z_2/t\|_2^2 \right)
\]

• $x_1$-update: linear equation with coefficient $I + A^T A$

• $(x_2, x_3)$-update: decoupled evaluations of $\text{prox}_{t^{-1}f_1}$ and $\text{prox}_{t^{-1}f_2}$
Outline

- Douglas-Rachford splitting method
- examples
- alternating direction method of multipliers
- image deblurring example
- convergence
Image blurring model

\[ b = K x_t + w \]

- \( x_t \) is unknown image
- \( b \) is observed (blurred and noisy) image; \( w \) is noise
- \( N \times N \)-images are stored in column-major order as vectors of length \( N^2 \)

Blurring matrix \( K \)

- represents 2D convolution with space-invariant point spread function
- with periodic boundary conditions, block-circulant with circulant blocks
- can be diagonalized by multiplication with unitary 2D DFT matrix \( W \):

\[ K = W^H \text{diag}(\lambda) W \]

equations with coefficient \( I + K^T K \) can be solved in \( O(N^2 \log N) \) time
Total variation deblurring with 1-norm

minimize $\|Kx - b\|_1 + \gamma\|Dx\|_{tv}$
subject to $0 \leq x \leq 1$

second term in objective is total variation penalty

$Dx$ is discretized first derivative in vertical and horizontal direction

$D = \begin{bmatrix} I \otimes D_1 \\ D_1 \otimes I \end{bmatrix}$, $D_1 = \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix}$

$\| \cdot \|_{tv}$ is a sum of Euclidean norms: $\|(u, v)\|_{tv} = \sum_{i=1}^{n} \sqrt{u_i^2 + v_i^2}$
Solution via Douglas-Rachford method

an example of a composite optimization problem

\[
\text{minimize } \ f_1(x) + f_2(Ax)
\]

with \( f_1 \) the indicator of \([0, 1]^n\) and

\[
A = \begin{bmatrix} K \\ D \end{bmatrix}, \quad f_2(u, v) = \|u\|_1 + \gamma\|v\|_{tv}
\]

**Primal DR method** (page 13-11) and **ADMM** (page 13-19) require:

- decoupled prox-evaluations of \( \|u\|_1 \) and \( \gamma\|v\|_{tv} \), and projections on \( C \)
- solution of linear equations with coefficient matrix

\[
I + K^T K + D^T D
\]

solvable in \( O(N^2 \log N) \) time
Example

- $1024 \times 1024$ image, periodic boundary conditions
- Gaussian blur
- salt-and-pepper noise (50% pixels randomly changed to 0/1)

original | noisy/blurred | restored
Convergence

cost per iteration is dominated by 2D FFTs

Douglas-Rachford method and ADMM
Outline

- Douglas-Rachford splitting method
- examples
- alternating direction method of multipliers
- image deblurring example
- convergence
Douglas-Rachford iteration mappings

define iteration map $F$ and negative step $G$ (in notation of page 13-7)

\[
F(y) = y + \text{prox}_g(2\text{prox}_f(y) - y) - \text{prox}_f(y)
\]
\[
G(y) = y - F(y)
\]
\[
= \text{prox}_f(y) - \text{prox}_g(2\text{prox}_f(y) - y)
\]

- $F$ is firmly nonexpansive (co-coercive with parameter 1)

\[
(F(y) - F(\hat{y}))^T(y - \hat{y}) \geq \|F(y) - F(\hat{y})\|_2^2 \quad \forall y, \hat{y}
\]

- this implies that $G$ is firmly nonexpansive:

\[
(G(y) - G(\hat{y}))^T(y - \hat{y})
\]
\[
= \|G(y) - G(\hat{y})\|_2^2 + (F(y) - F(\hat{y}))^T(y - \hat{y}) - \|F(y) - F(\hat{y})\|_2^2
\]
\[
\geq \|G(y) - G(\hat{y})\|_2^2
\]
Proof (of firm nonexpansiveness of $F$).

- define $x = \text{prox}_f(y)$, $\hat{x} = \text{prox}_f(\hat{y})$, and
  \[ v = \text{prox}_g(2x - y), \quad \hat{v} = \text{prox}_g(2\hat{x} - \hat{y}) \]

- substitute expressions $F(y) = y + v - x$ and $F(\hat{y}) = \hat{y} + \hat{v} - \hat{x}$:
  \[
  (F(y) - F(\hat{y}))^T (y - \hat{y}) \\
  \geq (y + v - x - \hat{y} - \hat{v} + \hat{x})^T (y - \hat{y}) - (x - \hat{x})^T (y - \hat{y}) + \|x - \hat{x}\|^2_2 \\
  = (v - \hat{v})^T (y - \hat{y}) + \|y - x - \hat{y} + \hat{x}\|^2_2 \\
  = (v - \hat{v})^T (2x - y - 2\hat{x} + \hat{y}) - \|v - \hat{v}\|^2_2 + \|F(y) - F(\hat{y})\|^2_2 \\
  \geq \|F(y) - F(\hat{y})\|^2_2
  
  inequalities use firm nonexpansiveness of $\text{prox}_f$ and $\text{prox}_g$ (page 6-9):
  \[
  (x - \hat{x})^T (y - \hat{y}) \geq \|x - \hat{x}\|^2_2, \quad (2x - y - 2\hat{x} + \hat{y})^T (v - \hat{v}) \geq \|v - \hat{v}\|^2_2
  \]
Convergence result

\[ y^{(k)} = (1 - \rho_k) y^{(k-1)} + \rho_k F(y^{(k-1)}) \]
\[ = y^{(k-1)} - \rho_k G(y^{(k-1)}) \]

Assumptions

- \( F \) has fixed points (points \( x \) that satisfy \( 0 \in \partial f(x) + \partial g(x) \))
- \( \rho_k \in [\rho_{\text{min}}, \rho_{\text{max}}] \) with \( 0 < \rho_{\text{min}} < \rho_{\text{max}} < 2 \)

Result

- \( y^{(k)} \) converges to a fixed point \( y^* \) of \( F \)
- \( x^{(k)} = \text{prox}_f(y^{(k-1)}) \) converges to a solution \( x^* = \text{prox}_f(y^*) \)
  (follows from continuity of \( \text{prox}_f \))
Proof: let \( y^* \) be any fixed point of \( F(y) \) (zero of \( G(y) \))

consider iteration \( k \) (with \( y = y^{(k-1)} \), \( \rho = \rho_k \), \( y^+ = y^{(k)} \)):

\[
\|y^+ - y^*\|_2^2 - \|y - y^*\|_2^2 = 2(y^+ - y)^T(y - y^*) + \|y^+ - y\|_2^2
\]
\[
= -2\rho G(y)^T(y - y^*) + \rho^2\|G(y)\|_2^2
\]
\[
\leq -\rho(2 - \rho)\|G(y)\|_2^2
\]
\[
\leq -M\|G(y)\|_2^2
\]  

(1)

where \( M = \rho_{\text{min}}(2 - \rho_{\text{max}}) \) (on line 3 we use firm nonexpansiveness of \( G' \))

• (1) implies that

\[
M \sum_{k=0}^{\infty} \|G(y^{(k)})\|_2^2 \leq \|y^{(0)} - y^*\|_2^2, \quad \|G(y^{(k)})\|_2 \to 0
\]

• (1) implies that \( \|y^{(k)} - y^*\|_2 \) is nonincreasing; hence \( y^{(k)} \) is bounded

• since \( \|y^{(k)} - y^*\|_2 \) is nonincreasing, the limit \( \lim_{k \to \infty} \|y^{(k)} - y^*\|_2 \) exists
Proof (continued)

• since the sequence \( y^{(k)} \) is bounded, it has a convergent subsequence

• let \( \bar{y}_k \) be a convergent subsequence with limit \( \bar{y} \); by continuity of \( G \),

\[
0 = \lim_{k \to \infty} G(\bar{y}_k) = G(\bar{y})
\]

hence, \( \bar{y} \) is a zero of \( G \) and the limit \( \lim_{k \to \infty} \| y^{(k)} - \bar{y} \|_2 \) exists

• let \( \bar{y}_1 \) and \( \bar{y}_2 \) be two limit points; the limits

\[
\lim_{k \to \infty} \| y^{(k)} - \bar{y}_1 \|_2, \quad \lim_{k \to \infty} \| y^{(k)} - \bar{y}_2 \|_2
\]

exist, and subsequences of \( y^{(k)} \) converge to \( \bar{y}_1 \), resp. \( \bar{y}_2 \); therefore

\[
\| \bar{y}_2 - \bar{y}_1 \|_2 = \lim_{k \to \infty} \| y^{(k)} - \bar{y}_1 \|_2 = \lim_{k \to \infty} \| y^{(k)} - \bar{y}_2 \|_2 = 0
\]
References


The image deblurring example is taken from this paper.