# LOCAL LINEAR CONVERGENCE OF ADMM

**Daniel Boley**

$$\textbf{Model QP/LP:}\ \min {}^1\!/\!_2 \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} \text{ s.t. } A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0, \tag{1}$$

$$\textbf{Lagrangian:}\ \mathcal{L}(\mathbf{x}, \mathbf{y}) = {}^1\!/\!_2 \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} - \mathbf{y}^T \mathbf{x} \ \text{ s.t. } A\mathbf{x} = \mathbf{b}, \tag{2}$$

where $\mathbf{y} \geq 0$ is the vector of Lagrange multipliers for the inequality constraints $\mathbf{x} \geq 0$.

# Applications

- Machine Learning
  "minimize some loss fcn subject to fitting training data"
- Economics, Operations Research, many others.

# Properties

- Models are very large
- Often subproblems can be solved easily
- Leads to idea of splitting problem into easy pieces.

# Previous Convergence Theory

- Very abstract theory based on monotone linear operators.
- Recent results are of the form $O(k)$ or $O(k^2)$, where $k$ = iteration number.
- Bounds are far from actual behavior.

# Dual Ascent Method

**Model QP/LP:** $\min {}^1\!/\!_2\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x}$ s.t. $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq 0$, (1)

**Lagrangian:** $\mathcal{L}(\mathbf{x}, \mathbf{y}) = {}^1\!/\!_2\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} - \mathbf{y}^T\mathbf{x}$ s.t. $A\mathbf{x} = \mathbf{b}$, (2)

where $\mathbf{y} \geq 0$ = Lagrange multipliers for the constraints $\mathbf{x} \geq 0$.

Primal Problem: $\min_{\mathbf{x}} \boxed{\max_{\mathbf{y}} \mathcal{L}(\mathbf{x}, \mathbf{y})}$: $\boxed{\cdots} = \infty$ when constraints violated.

Dual Problem: $\max_{\mathbf{y}} \boxed{\min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{y})}$: $\boxed{\textit{boxed expr}}$ is relatively easy to solve.

Dual Ascent Method: solve $\boxed{\min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{y})}$ in dual problem exactly, take small gradient ascent steps on dual variable $\mathbf{y}$.

## Split Primal variables into x, z:

$$\min {}^1\!/_2 \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} + g(\mathbf{z}) \text{ s.t. } A\mathbf{x} = \mathbf{b},\ \mathbf{x} = \mathbf{z}, \tag{3}$$

where $g(\mathbf{z})$ is the indicator function for the non-negative orthant:

$$g(\mathbf{z}) = \begin{cases} 0 & \text{if } \mathbf{z} \geq 0 \\ \infty & \text{if any component of } \mathbf{z} \text{ is negative.} \end{cases}.$$

$g(\mathbf{z})$ is a non-smooth convex function encoding the inequality constraints.

## Associated [partially] augmented Lagrangian

$$\mathcal{L}_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}) = {}^1\!/_2 \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} + g(\mathbf{z}) + \mathbf{y}^T(\mathbf{x} - \mathbf{z}) + {}^1\!/_2 \rho \|\mathbf{x} - \mathbf{z}\|_2^2, \quad \text{s.t. } A\mathbf{x} = \mathbf{b}, \tag{4}$$

where $\mathbf{y}$ is the vector of Lagrange multipliers for the additional equality constraint $\mathbf{x} - \mathbf{z} = 0$, $\rho$ is a proximity penalty parameter chosen by the user.

# Splitting

Using the common splitting boyd11, the ADMM method consists of three steps: first minimize Lagrangian with respect to $\mathbf{x}$, then with respect to $\mathbf{z}$, and then perform one ascent step on the Lagrange multipliers $\mathbf{u}$:

$$\begin{array}{lll}
1. \text{ Set } \mathbf{x}^{[k+1]} & = & \operatorname{argmin}_{\mathbf{x}} {}^1\!/_2\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + {}^1\!/_2\rho\mathbf{x}^T\mathbf{x} + \rho\mathbf{x}^T(\mathbf{u}^{[k]} - \mathbf{z}^{[k]}) \\
 & & \text{subject to } A\mathbf{x} = \mathbf{b} \\
2. \text{ Set } \mathbf{z}^{[k+1]} & = & \operatorname{argmin}_{\mathbf{z}} g(\mathbf{z}) + {}^1\!/_2\rho\mathbf{z}^T\mathbf{z} - \rho\mathbf{z}^T(\mathbf{x}^{[k+1]} + \mathbf{u}^{[k]}) \\
3. \text{ Set } \mathbf{u}^{[k+1]} & = & \mathbf{u}^{[k]} + \nabla_{\mathbf{u}}\mathcal{L}_\rho(\mathbf{x}^{[k+1]}, \mathbf{z}^{[k+1]}, \mathbf{u}).
\end{array} \tag{5}$$

# Closed Form

Each step of Alg I can be solved in closed form, leading to the ADMM iteration (with no acceleration) consisting of the following steps repeated until convergence, where $\mathbf{z}^{[k]}, \mathbf{u}^{[k]}$ denote the vectors from the previous pass, and $\rho$ is a given fixed proximity penalty:

---

Algorithm 1: One Pass of ADMM

---

Start with $\mathbf{z}^{[k]}, \mathbf{u}^{[k]}$.

1. Solve $\begin{pmatrix} Q + \rho I & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}^{[k+1]} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \rho(\mathbf{z}^{[k]} - \mathbf{u}^{[k]}) - \mathbf{c} \\ \mathbf{b} \end{pmatrix}$ for $\mathbf{x}^{[k+1]}, \boldsymbol{\nu}$.
2. Set $\mathbf{z}^{[k+1]} = \max\{0, \mathbf{x}^{[k+1]} + \mathbf{u}^{[k]}\}$ (where "max" is taken elementwise).
3. Set $\mathbf{u}^{[k+1]} = \mathbf{u}^{[k]} + \mathbf{x}^{[k+1]} - \mathbf{z}^{[k+1]}$.

Result is $\mathbf{z}^{[k+1]}, \mathbf{u}^{[k+1]}$ for next pass.

---

# Complementarity Property

**Lemma 1.** *After every pass, the vectors* $\mathbf{z}^{[k+1]}, \mathbf{u}^{[k+1]}$ *satisfy*

*a.* $\mathbf{z}^{[k+1]} \geq 0$,

*b.* $\mathbf{u}^{[k+1]} \leq 0$,

*c.* $z_i^{[k+1]} \cdot u_i^{[k+1]} = 0, \forall i$ *(a complementarity condition).*

*d.* $\mathbf{x}^{[k+1]}$ *satisfies the equality constraints* $A\mathbf{x}^{[k+1]} = \mathbf{b}$.

# Combine Iterates into a Single Vector

Use the complementarity condition to store $\mathbf{z}, \mathbf{u}$ in a single vector.

Let $\mathbf{w} = \mathbf{z} - \mathbf{u}$, and let $\mathbf{d}$ be a vector of flags such that

$$\begin{aligned} d_i &= +1 \quad \textit{iff} \quad u_i = 0 \quad \Longleftrightarrow \quad i\text{-th constraint is inactive}, \\ d_i &= -1 \quad \textit{iff} \quad u_i \neq 0 \quad \Longleftrightarrow \quad i\text{-th constraint is active}. \end{aligned}$$

Because of the complementarity condition, $z_i = {}^1\!/_2(1 + d_i)w_i$ and $u_i = -{}^1\!/_2(1 - d_i)w_i$. If $D = \textsc{Diag}(\mathbf{d})$ (the diagonal matrix with elements of vector $\mathbf{d}$ on the diagonal), then ${}^1\!/_2(I-D)\mathbf{w} = -\mathbf{u}$ and ${}^1\!/_2(I+D)\mathbf{w} = \mathbf{z}$. The flags indicate which inequality constraints are actively enforced on $\mathbf{z}$ at each pass.

# Single Vector Iteration

The modified ADMM iteration using the new variables can be written as:

---

Algorithm 2: One Pass of Modified ADMM

---

Start with $\mathbf{w}^{[k]}, D^{[k]}$.

1. Solve $\begin{pmatrix} Q/\rho + I & A^T/\rho \\ A & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}^{[k+1]} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^{[k]} - \mathbf{c}/\rho \\ \mathbf{b} \end{pmatrix}$
   for $\mathbf{x}^{[k+1]}, \boldsymbol{\nu}$.
2. Set $\mathbf{w}_{\text{tmp}} = \mathbf{x}^{[k+1]} - {}^1\!/\!_2(I - D^{[k]})\mathbf{w}^{[k]}$, where $D^{[k]} = \textsc{Diag}(\mathbf{d}^{[k]})$;
3. $D^{[k+1]} = \textsc{Diag}(\textsc{sign}(\mathbf{w}_{\text{tmp}}))$;
4. Set $\mathbf{w}^{[k+1]} = |\mathbf{w}_{\text{tmp}}| = D^{[k+1]}\mathbf{w}_{\text{tmp}}$.

Result is $\mathbf{w}^{[k+1]}, D^{[k+1]}$ for next pass.

---

# Eliminate x

Iteration is carried by $\mathbf{z}, \mathbf{u}$. So eliminate $\mathbf{x}$ entirely, by inverting the matrix in step one explicitly.

$$\begin{aligned} \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\nu} \end{pmatrix} &= \begin{pmatrix} Q/\rho + I & A^T/\rho \\ A & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{w} - \mathbf{c}/\rho \\ \mathbf{b} \end{pmatrix} \\ &= \begin{pmatrix} N & RA^TS \\ \rho SAR & -\rho S \end{pmatrix} \begin{pmatrix} \mathbf{w} - \mathbf{c}/\rho \\ \mathbf{b} \end{pmatrix}, \end{aligned} \tag{6}$$

where $R = (Q/\rho + I)^{-1}$ is the resolvent of $Q$, $S = (ARA^T)^{-1}$ is the inverse of the Schur complement, and $N = R - RA^TSAR$.

# Matrix Form of Iteration

**Lemma 2.** *The operator $N = R - RA^TSAR$ is positive semi-definite and $\|N\|_2 \leq \|R\|_2 \leq 1$. If $Q$ is strictly positive definite, then also $\|R\|_2 < 1$. If we have an LP, then $N = I - A^T(AA^T)^{-1}A$ = orthogonal projector onto the nullspace of $A$.*

# ADMM as a Matrix Recurrence

Combine formulas

$$\mathbf{x}^{[k+1]} = N\mathbf{w}^{[k]} + \overbrace{RA^T S\mathbf{b} - N\mathbf{c}/\rho}^{\mathbf{h}}$$

$$\mathbf{w}_{\text{tmp}} = \mathbf{x}^{[k+1]} - {}^1\!/_2(I - D^{[k]})\mathbf{w}^{[k]}$$

$$D^{[k+1]} = \textsc{Diag}(\textsc{sign}(\mathbf{w}_{\text{tmp}}))$$

$$\mathbf{w}^{[k+1]} = |\mathbf{w}_{\text{tmp}}| = D^{[k+1]}\mathbf{w}_{\text{tmp}}$$

to get

---

Algorithm 3: One Pass of Reduced ADMM

---

Start with $\mathbf{w}^{[k]}, D^{[k]}$.

0. $\mathbf{w}_{\text{tmp}} \quad = \quad (N - {}^1\!/_2(I - D^{[k]}))\mathbf{w}^{[k]} + \mathbf{h}$
1. $D^{[k+1]} \quad = \quad \textsc{Diag}(\textsc{sign}(\mathbf{w}_{\text{tmp}}))$
2. $\mathbf{w}^{[k+1]} \quad = \quad D^{[k+1]}\mathbf{w}_{\text{tmp}}$

Result is $\mathbf{w}^{[k+1]}, D^{[k+1]}$ for next pass.

---

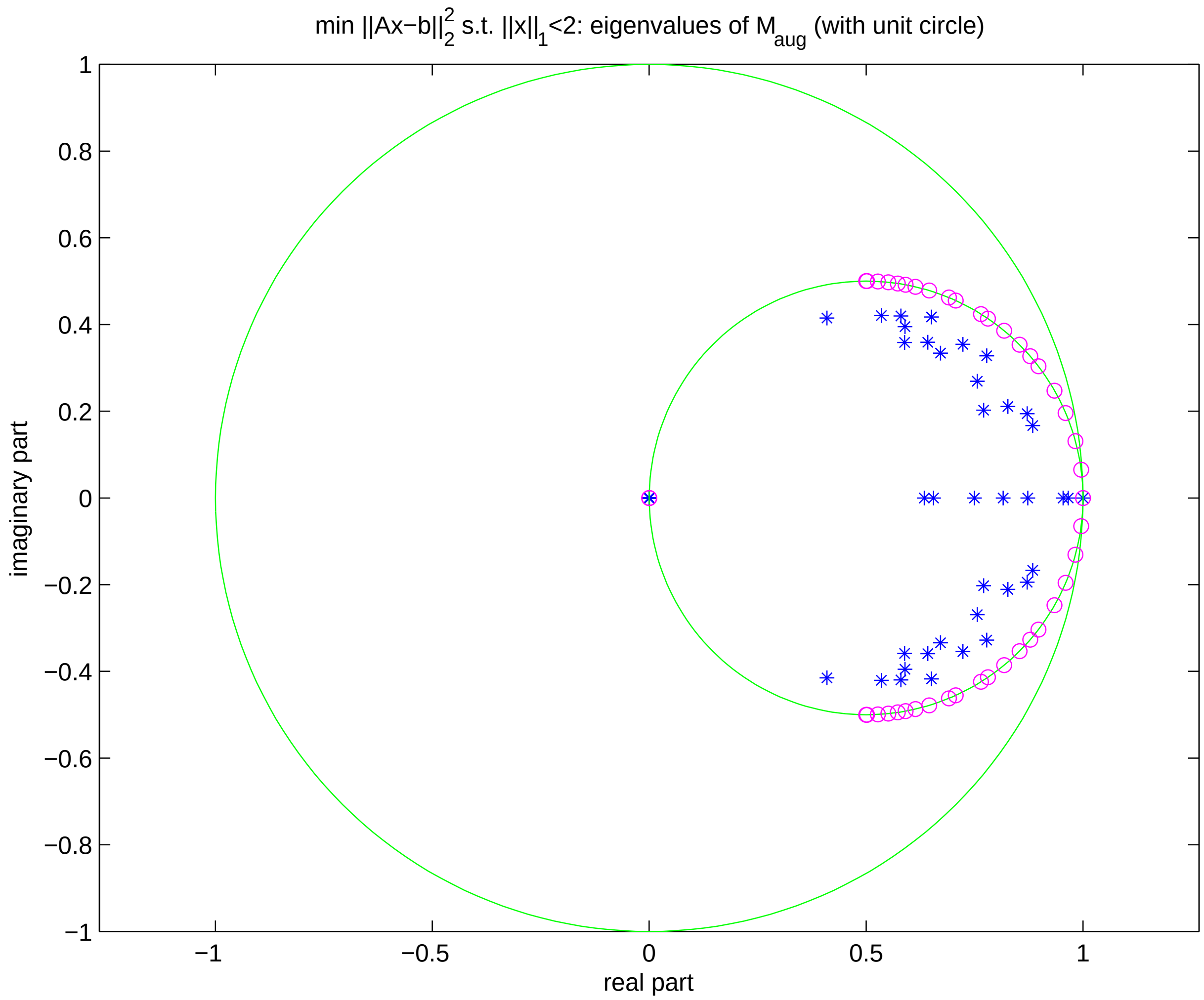
# Spectral Properties

The spectral properties of $M^{[k]} = D^{[k+1]}(N - {}^1\!/_2(I - D^{[k]}))$ play a critical role in the convergence of this procedure.

**Lemma 3.** $\|M\|_2 = \|D^{[k+1]}(N - {}^1\!/_2(I - D^{[k]}))\|_2 \le 1$. *Any eigenvalues of* $M = D^{[k+1]}(N - {}^1\!/_2(I - D^{[k]}))$ *on the unit circle must have a complete set of eigenvectors (no Jordan blocks larger than* $1 \times 1$*).*

**Lemma 4.** *If* $D = D^{[k+1]} = D^{[k]}$ *(the flags remain unchanged), then all eigenvalues of* $D(N - {}^1\!/_2(I - D))$ *must lie in the closed disk in the complex plane with center* ${}^1\!/_2$ *and radius* ${}^1\!/_2$*, denoted* $\mathcal{D}({}^1\!/_2, {}^1\!/_2)$*. The only possible eigenvalue on the unit circle is* $+1$*, and if present must have a complete set of eigenvectors. In the case of a linear program,* $Q = 0$*,* $N$ *is an orthogonal projector, and all the eigenvalues of* $M = D(N - {}^1\!/_2(I - D))$ *lie on the boundary of* $\mathcal{D}({}^1\!/_2, {}^1\!/_2)$*.*

# Example: Spectrum of ADMM Iteration Operator




○ = eigenvalues for LP near end of iteration.

∗ = eigenvalues for QP.

# Matrix Recurrence.

Step 2 of Algorithm 3 is written as follows:

$$\begin{aligned} \begin{pmatrix} \mathbf{w}^{[k+1]} \\ 1 \end{pmatrix} &= \mathbf{M}_{\mathrm{aug}}^{[k]} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} = \begin{pmatrix} M^{[k]} & D^{[k+1]}\mathbf{h} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} D^{[k+1]}(N - {}^1\!/\!_2(I - D^{[k]})) & D^{[k+1]}\mathbf{h} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix}, \end{aligned} \tag{7}$$

where $\mathbf{h} = RA^T S\mathbf{b} - N\mathbf{c}/\rho$

Converges to eigenvector: if eigenvector is all non-negative, get solution to original QP/LP. Otherwise, the flag matrix ($D$) will change to yield a new operator.

# Solution to QP/LP is an eigenproblem.

**Lemma 5.** *Let* $\mathbf{M}_{\text{aug}}$ *be the augmented matrix in recurrence and assume* $D = D^{[k+1]} = D^{[k]}$ *is a flag matrix of the form* $\textsc{Diag}(\pm 1, \ldots, \pm 1)$*. Suppose* $\begin{pmatrix} \mathbf{w} \\ 1 \end{pmatrix}$ *is an eigenvector corresponding to eigenvalue 1 of the matrix* $\mathbf{M}_{\text{aug}}$ *and furthermore suppose* $\mathbf{w} \geq 0$*.*

**Then** *the primal variables defined by* $\mathbf{x} = \mathbf{z} = {}^1\!/_2 (I+D)\mathbf{w}$ *and dual variables* $\mathbf{y} = \rho\mathbf{u} = -{}^{\rho}\!/_2 (I-D)\mathbf{w}$ *satisfy the first order KKT conditions.*

<u>**Conversely,**</u> *if* $\mathbf{x} = \mathbf{z}, \mathbf{u}$ *satisfy the KKT conditions,*

**then** $\begin{pmatrix} \mathbf{w} \\ 1 \end{pmatrix}$ *is an eigenvector of* $\mathbf{M}_{\text{aug}}$ *corresponding to eigenvalue 1, where* $\mathbf{w} = \mathbf{z} - \mathbf{u}$, *and* $\mathbf{M}_{\text{aug}}$ *is defined as in the recurrence with* $D^{[k+1]} = D^{[k]} = D = \textsc{Diag}(\mathbf{d})$ *with entries* $d_i = +1$ *if* $z_i > 0$, $d_i = -1$ *if* $u_i < 0$, *else* $d_i = \pm 1$ *(either sign).*

## Regimes based on spectral properties.

If $D^{[k+1]} = D^{[k]}$:

[a] The spectral radius of $M^{[k]}$ is strictly less than 1. If close enough to the optimal solution (if it exists), the result is linear convergence to that solution.

[b] $M^{[k]}$ has an eigenvalue equal to 1 which results in a $2 \times 2$ Jordan block for $\mathbf{M}_{\text{aug}}^{[k]}$. The process tends to a constant step, either diverging, or driving some component negative, resulting in a change in the operator $M^{[k]}$.

[c] $M^{[k]}$ has an eigenvalue equal to 1, but $\mathbf{M}_{\text{aug}}^{[k]}$ still has no non-diagonal Jordan block for eigenvalue 1; If close enough to the optimal solution (if it exists), the result is linear convergence to that solution.

If $D^{[k+1]} \neq D^{[k]}$, then we transition to a new operator:

[d] $M^{[k]}$ has have an eigenvalue of absolute value 1, but not equal to 1. This can occur when the iteration transitions to a new set of active constraints.

# Local Convergence is Linear.

**Theorem 6.** *Suppose the LP/QP has a unique solution* $\mathbf{x}^* = \mathbf{z}^*$ *and corresponding unique optimal Lagrange multipliers* $\mathbf{y}^*$ *for the inequality constraints, and this solution has strict complementarity: that is either* $z_i^* > 0 = y_i^*$ *or* $y_i^* < 0 = z_i^*$ *(i.e.* $w_i^* = z_i^* - y_i^*/\rho > 0$*) for every index* $i$*. Then eventually the ADMM iteration reaches a stage where it converges linearly to that unique solution.*

## Example: A Simple Basis Pursuit Problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } A\mathbf{x} = \mathbf{b}, \tag{8}$$

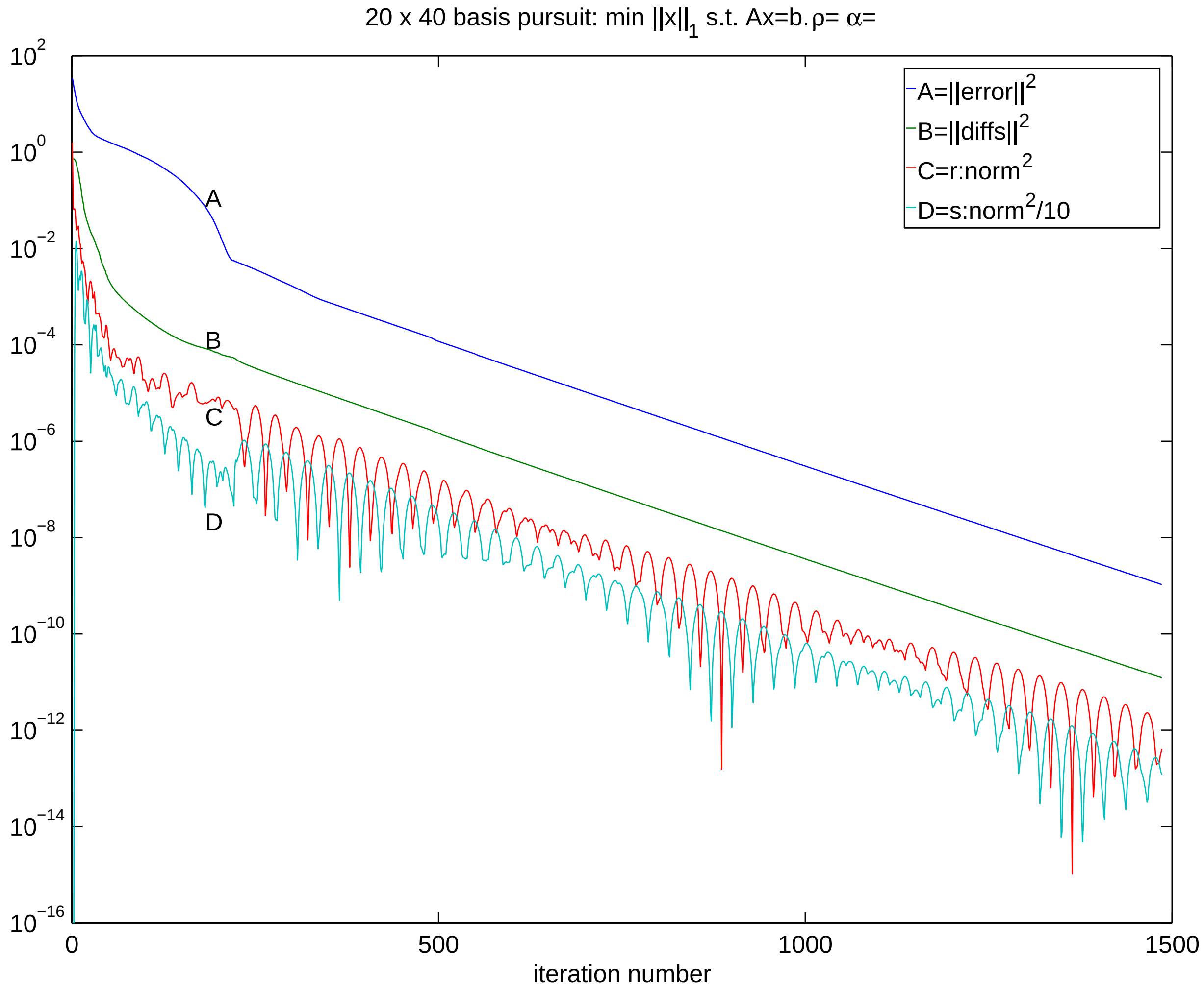
or a soft variation allowing for noise (similar to LASSO )

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2^2 \text{ subject to } \|\mathbf{x}\|_1 \leq \text{tol}, \tag{9}$$

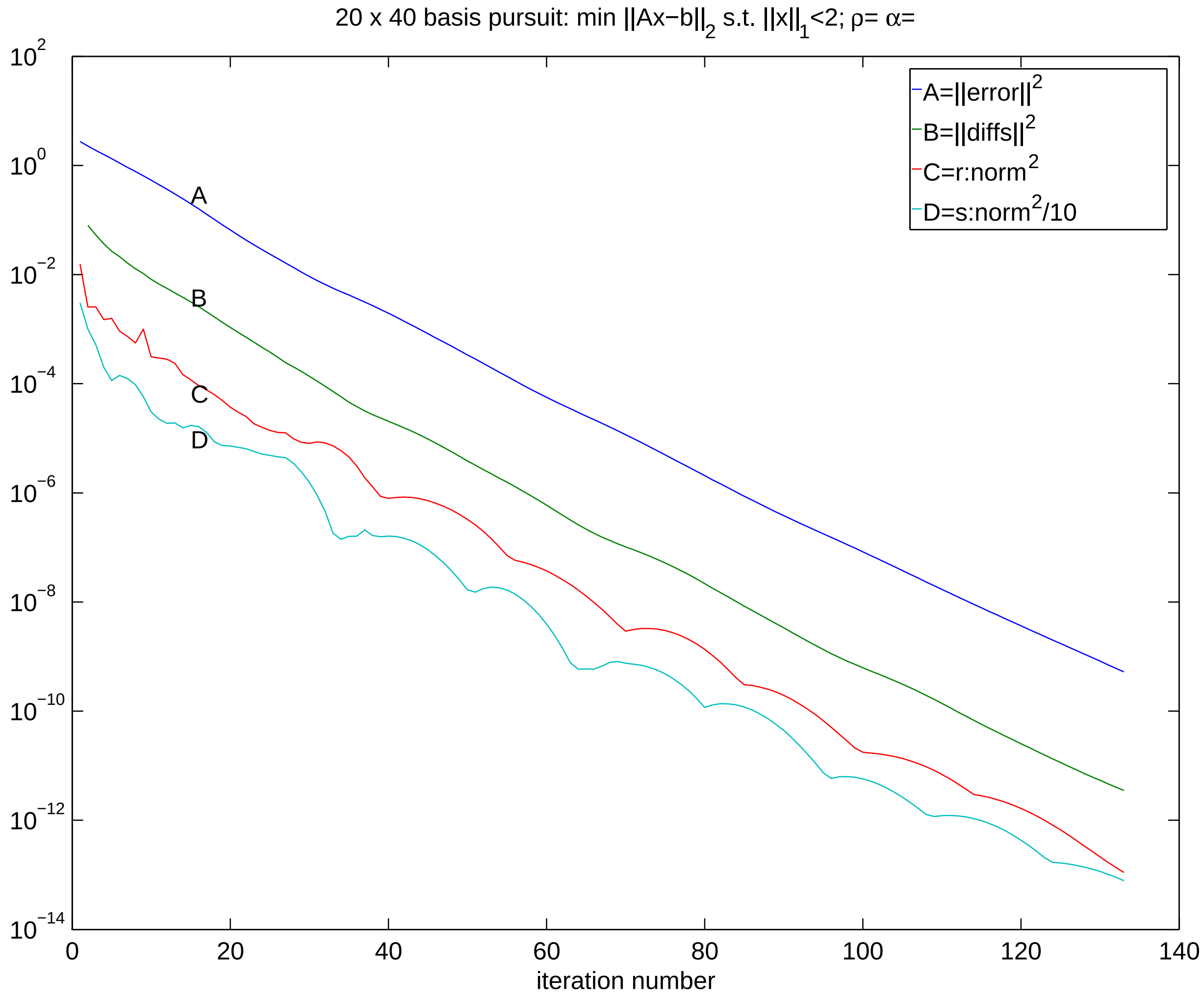
where the elements of $A, \mathbf{b}$ are generated independently by a uniform distribution over $[-1, +1]$. $A$ is $20 \times 40$.

Problem (9) is a model to find a sparse best fit, with a trade-off between goodness of fit and sparsity.

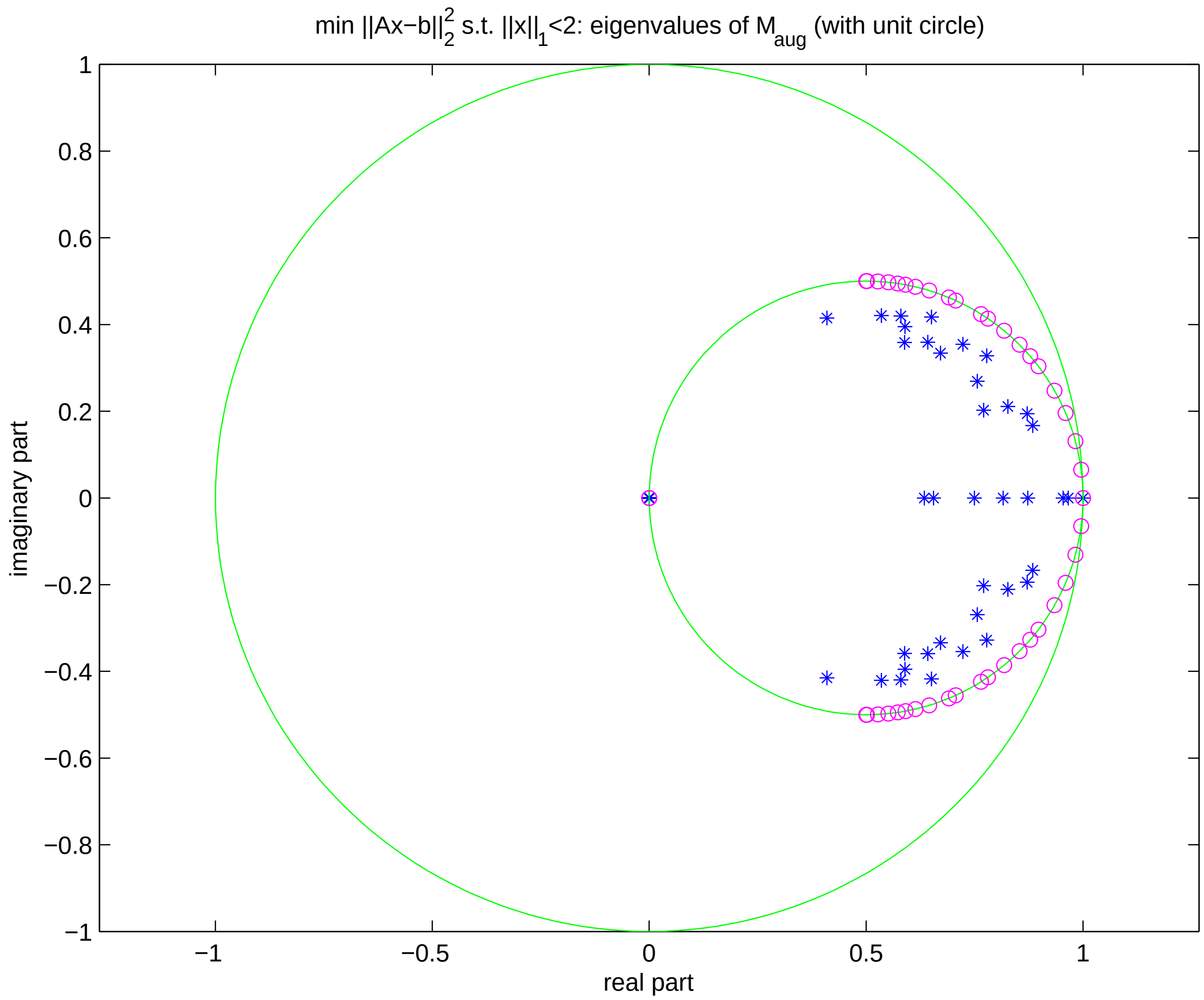
# ADMM applied to the Basis Pursuit LP



# Unaccelerated ADMM applied to the LASSO QP



# Spectrum of ADMM Iteration Operator – LASSO




○ = eigenvalues for LP in final regime.

$*$ = eigenvalues for QP.

# Toy Example

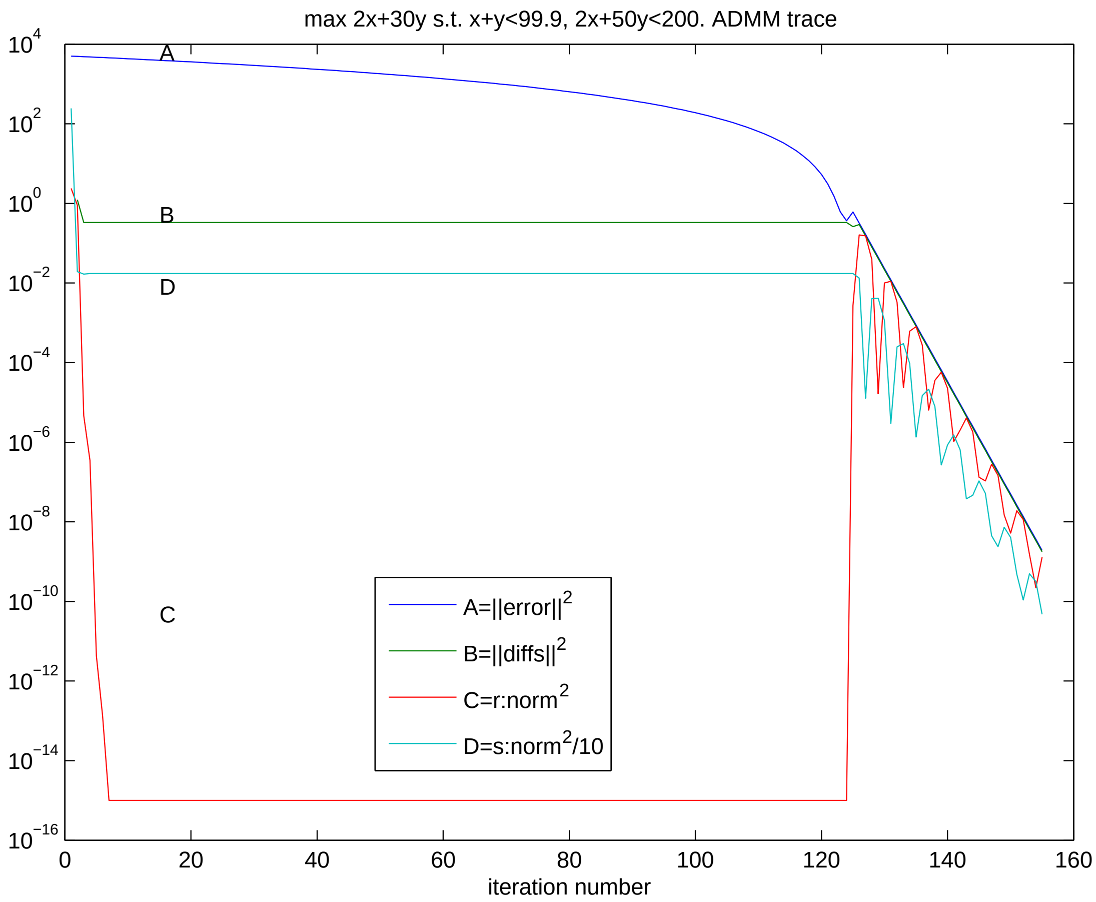
Simple resource allocation model:
- $x_1$ = rate of cheap process (e.g. fermentation),
- $x_2$ = rate of costly process (e.g. respiration).

$$
\begin{array}{llllll}
\text{maximize}_{\mathbf{x}} & +2x_1 + 30x_2 & & & & (\textit{desired end product production}) \\
\text{subject to} & x_1 \quad + x_2 & & \leq & x_{0,max} & (\textit{limit on raw material}) \\
& 2x_1 \quad + 50x_2 & & \leq & 200 & (\textit{internal capacity limit}) \\
& x_1 \geq 0 & x_2 \geq 0 & & & (\textit{irreversibility of reactions})
\end{array}
$$

Put into standard form:

$$
\begin{array}{llllll}
\text{minimize}_{\mathbf{x}} & -2x_1 - 30x_2 & & & & (\textit{desired end product production}) \\
\text{subject to} & x_1 \quad + x_2 & + \; x_3 & = & x_{0,max} & (\textit{limit on raw material}) \\
& 2x_1 \quad + 50x_2 & + \; x_4 & = & 200 & (\textit{internal capacity limit}) \\
& x_1 \geq 0 & x_2 \geq 0 & & & (\textit{irreversibility of reactions}) \\
& x_3 \geq 0 & x_4 \geq 0 & & & (\textit{slack variables})
\end{array}
\tag{10}
$$

# Typical Convergence Behavior $v_{0,max} = 99.9$




ADMM on Example 1: typical behavior. Curves: `A`: error $\|(\mathbf{z}^{[k]} - \mathbf{u}^{[k]}) - (\mathbf{z}^* - \mathbf{u}^*)\|^2$. `B`: $\|(\mathbf{z}^{[k]} - \mathbf{u}^{[k]}) - (\mathbf{z}^{[k-1]} - \mathbf{u}^{[k-1]})\|^2$. `C`: $\|(\mathbf{x}^{[k]} - \mathbf{z}^{[k]})\|^2$. `D`: $\|(\mathbf{z}^{[k]} - \mathbf{z}^{[k-1]})\|^2/10$ (`D` is scaled by $1/10$ just to separate it from the rest).

# Matrix Operators

$$N = \begin{pmatrix} 0.5201 & -0.0210 & -0.4991 & 0.0096 \\ -0.0210 & 0.0012 & 0.0197 & -0.0204 \\ -0.4991 & 0.0197 & 0.4793 & 0.0108 \\ 0.0096 & -0.0204 & 0.0108 & 0.9994 \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} 48.3546 \\ 2.0968 \\ 49.4487 \\ -1.5470 \end{pmatrix}.$$

ADMM Iterates for $k = 1, \ldots, 124$ follow:

$$\begin{pmatrix} \mathbf{w}^{[k+1]} \\ 1 \end{pmatrix} = \mathbf{M}_{\text{aug}} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} = \left(\begin{array}{cccc|c} 0.5201 & -0.0210 & -0.4991 & 0.0096 & 48.3546 \\ -0.0210 & 0.0012 & 0.0197 & -0.0204 & 2.0968 \\ -0.4991 & 0.0197 & 0.4793 & 0.0108 & 49.4487 \\ -0.0096 & 0.0204 & -0.0108 & 0.0006 & 1.5470 \\ \hline 0 & 0 & 0 & 0 & 1.0000 \end{array}\right) \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix}.$$

# Eigenstructure of Operator

The eigenvalues of the operator $\mathbf{M}_{\text{aug}}$ are given by its Jordan canonical form:

$$\mathbf{J} = \textsc{Diag}(\mathbf{J}_1, \, \mathbf{J}_4) = \textsc{Diag}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \, 6.2357\text{e-}4 \pm 2.4964\text{e-}2i, \, 0\right)$$

The $2 \times 2$ Jordan block corresponding to eigenvalue 1 indicates we are in the "constant-step" regime [b]. The difference between two consecutive iterates quickly converges to $\mathbf{M}_{\text{aug}}$'s only eigenvector for eigenvalue 1:

$$\begin{pmatrix} \mathbf{w}^{[k+1]} \\ 1 \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} \Longrightarrow \begin{pmatrix} 0.4160 \\ -0.0166 \\ -0.3993 \\ 0 \\ 0 \end{pmatrix},$$

# Final Regime

for $k = 133, \ldots, 154$:

$$\begin{pmatrix} \mathbf{w}^{[k+1]} \\ 1 \end{pmatrix} = \mathbf{M}_{\text{aug}} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} = \left(\begin{array}{rrrr|r} 0.5201 & -0.0210 & -0.4991 & 0.0096 & 48.3546 \\ -0.0210 & 0.0012 & 0.0197 & -0.0204 & 2.0968 \\ 0.4991 & -0.0197 & 0.5207 & -0.0108 & -49.4487 \\ -0.0096 & 0.0204 & -0.0108 & 0.0006 & 1.5470 \\ \hline 0 & 0 & 0 & 0 & 1.0000 \end{array}\right) \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix}, \tag{11}$$

# Final iterate:

$$\begin{pmatrix} \mathbf{w}^* \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{w}^{[155]} \\ 1 \end{pmatrix} = (99.8958,\ 0.0042,\ 0.8334,\ 0.5833,\ 1)^T\ .$$

(Eigenvector for $M$).

The final flag matrix is $D^* = \textsc{Diag}(+1, +1, -1, -1)$, indicating that the first two components of $\mathbf{w}^*$ correspond to primal variables $(x_1^*, x_2^*)$ and the last two to dual variables $(u_3^*, u_4^*)$, all non-zero.

Thus the true optimal solution to LP is $x_1^* = 99.8958$, $x_2^* = 0.0042$. $u_3^* = -0.8334$, $u_4^* = -0.5833$.

$\sigma(M) = 0.7217 \Longrightarrow$ fast convergence.

# Second Toy Example $v_{0,max} = 3.99$

For $k < 561$ in "constant step" regime.

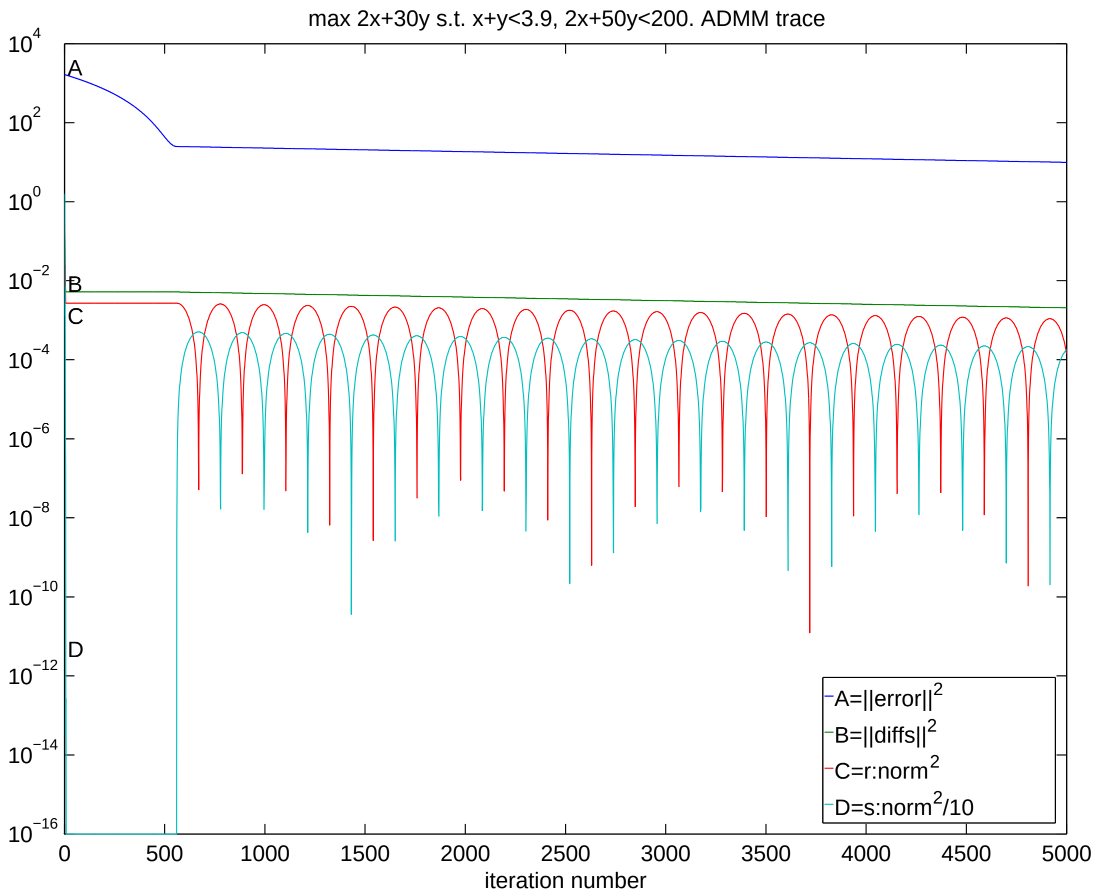
For $k \geq 561$:

$$\begin{pmatrix} \mathbf{w}^{[k+1]} \\ 1 \end{pmatrix} = \mathbf{M}_{\text{aug}} \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix} = \left(\begin{array}{rrrr|r} 0.4799 & 0.0210 & 0.4991 & -0.0096 & -0.4444 \\ -0.0210 & 0.0012 & 0.0197 & -0.0204 & 3.9924 \\ 0.4991 & -0.0197 & 0.5207 & -0.0108 & 0.5368 \\ 0.0096 & -0.0204 & 0.0108 & 0.9994 & -0.5094 \\ \hline 0 & 0 & 0 & 0 & 1.0000 \end{array}\right) \begin{pmatrix} \mathbf{w}^{[k]} \\ 1 \end{pmatrix}, \tag{12}$$

converging to eigenvector

$$\begin{pmatrix} \mathbf{w}^{*} \\ 1 \end{pmatrix} = \begin{pmatrix} 28.0 \\ 3.9 \\ 30.0 \\ 5.0 \\ 1 \end{pmatrix}.$$

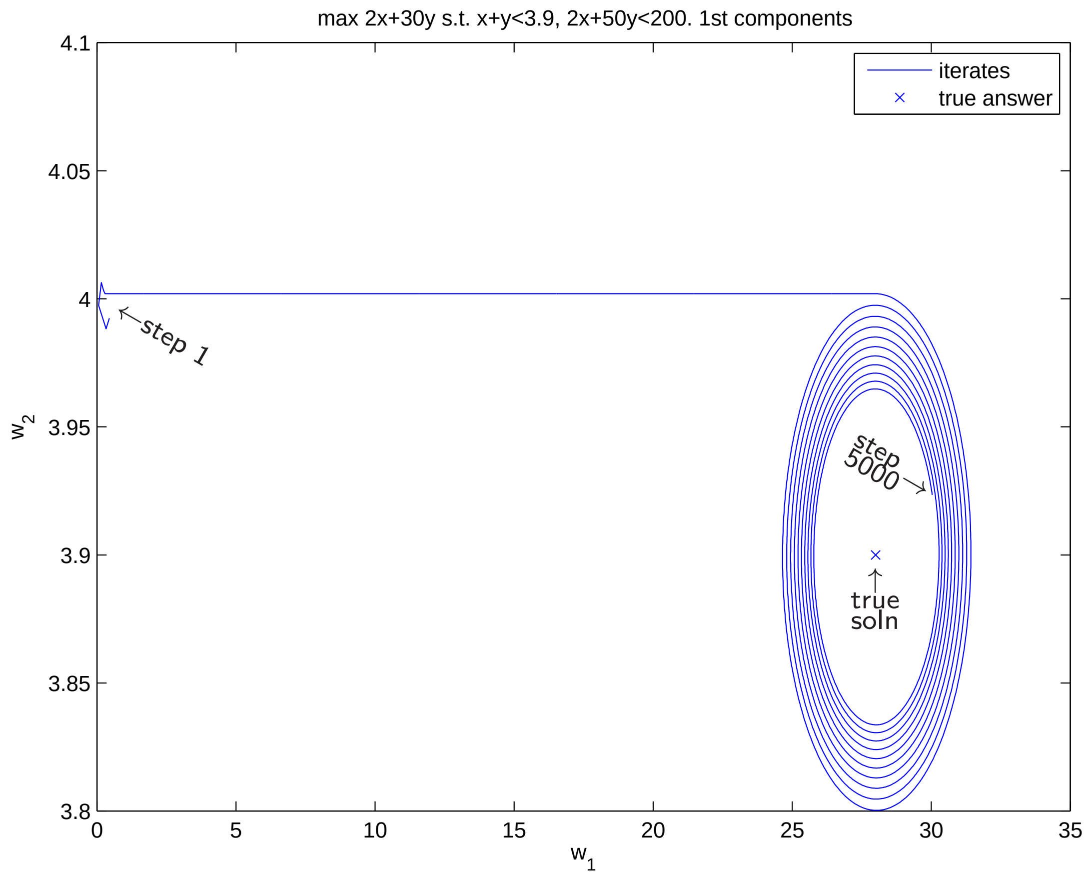
# Convergence Of Second Toy Example




ADMM on Example 2: slow linear convergence.

Second largest eigenvalue $= \sigma(M) = 0.999896$. convergence is very slow: $-1/\log_{10}(\sigma(M)) = 22135$ iterations needed per decimal digit of accuracy.

# Convergence Of Second Toy Example




Convergence behavior of first two components of $\mathbf{w}^{[k]}$ for Example 2, showing the initial straight line behavior (initial regime [b]) leading to the spiral (final regime [a]).

# FUTURE WORK

ADMM $\Longleftrightarrow$ power method with different operators, changing with regime. Replace power method with faster eigensolver.

Conduct similar analysis on other patterns (e.g. LASSO).

Discover relation between eigenvalues controlling convgernce rate and original QP/LP.