INTERPRETING DUAL VARIABLES: DUAL OF THE DIET PROBLEM

Diet Problem Data

$$F : \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rp} \end{bmatrix}$$

FOODS

NUTRIENTS

$$\ell : \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_p \end{pmatrix} \quad \mathbf{u} : \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$$

Diet Problem Data

Food	Calories	Total_Fat	Protein	Vit_A	Vit_C	Calcium	Price
Peppers	20	0.1	0.7	467.7	66.1	6.7	0.8
Potatoes, Baked	171.5	0.2	3.7	0	15.6	22.7	0.5
Tofu	88.2	5.5	9.4	98.6	0.1	121.8	1.1
Couscous	100.8	0.1	3.4	0	0	7.2	1
White Rice	102.7	0.2	2.1	0	0	7.9	0.4
Macaroni,Ckd	98.7	0.5	3.3	0	0	4.9	0.2
Peanut Butter	188.5	16	7.7	0	0	13.1	0.6

Nutrient	Min	Max
Calories	2000	2250
Total_Fat	0	65
Protein	50	100
Vit A	5000	50000
Vit C	50	20000
Calcium	800	1600

Diet Problem Setup

$$\begin{array}{rcccl} \min & & \mathbf{c}^\intercal \mathbf{x} & & \\ \ell & \leq & F^\intercal \mathbf{x} & \leq & \mathbf{u} \\ & \mathbf{x} & \geq & \mathbf{0} & \end{array}$$

$$\boxed{\begin{array}{lrcl} \max & -\mathbf{c}^\intercal \mathbf{x} & & \\ & F^\intercal \mathbf{x} & \leq & \mathbf{u} \\ & -F^\intercal \mathbf{x} & \leq & -\ell \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}}$$

Diet Problem Dual

$$\begin{array}{lrcl} \max & -\mathbf{c}^\intercal\ \mathbf{x} & & \\ & F^\intercal \mathbf{x} & \leq & \mathbf{u} \\ & -F^\intercal \mathbf{x} & \leq & -\ell \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

$$\begin{array}{lrcl} \min & \mathbf{u}^\intercal \mathbf{y_u} - \ell^\intercal\ \mathbf{y}_\ell & & \\ & F(\mathbf{y_u} - \mathbf{y}_\ell) & \geq & -\mathbf{c} \\ & \mathbf{y_u},\ \mathbf{y}_\ell & \geq & \mathbf{0} \end{array}$$

What does the dual mean?

Food	Calories	Total_Fat	Protein	Vit_A	Vit_C	Calcium	Price
Peppers	20	0.1	0.7	467.7	66.1	6.7	0.8
Potatoes, Baked	171.5	0.2	3.7	0	15.6	22.7	0.5
Tofu	88.2	5.5	9.4	98.6	0.1	121.8	1.1
Couscous	100.8	0.1	3.4	0	0	7.2	1
White Rice	102.7	0.2	2.1	0	0	7.9	0.4
Macaroni,Ckd	98.7	0.5	3.3	0	0	4.9	0.2
Peanut Butter	188.5	16	7.7	0	0	13.1	0.6

Nutrient	Min	Max
Calories	2000	2250
Total_Fat	0	65
Protein	50	100
Vit A	5000	50000
Vit C	50	20000
Calcium	800	1600

Optimal Solutions

Food	Opt. Amt.
Peppers	9.55
Potatoes, Baked	0.95
Tofu	5.39
Couscous	0.00
White Rice	0.00
Macaroni,Ckd	11.86
Peanut Butter	0.00

Nutrient	Dual (yU)	Dual (yL)
Calories	0.000	0.002
Total_Fat	0.000	0.000
Protein	0.021	0.000
Vit A	0.000	0.002
Vit C	0.000	0.000
Calcium	0.000	0.008

SHADOW COSTS

sensitivity analysis

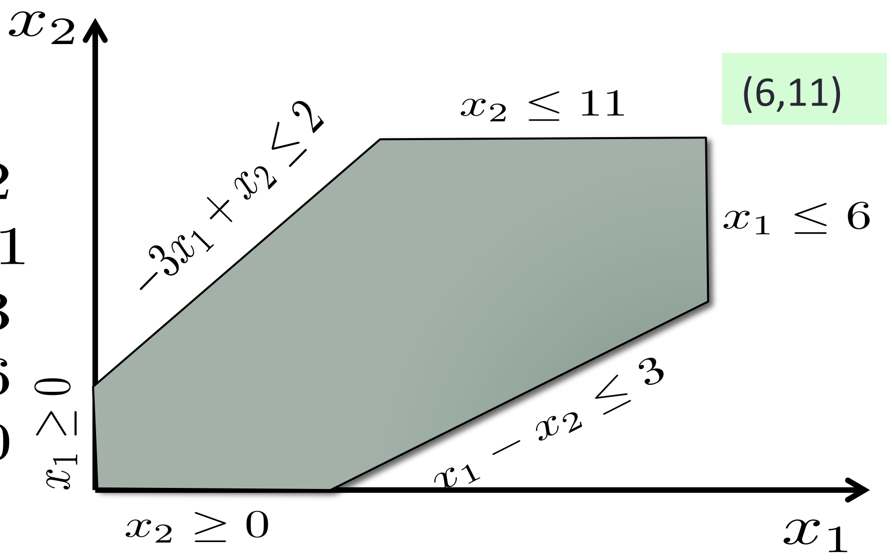
Shadow Cost of Constraints

$$\begin{array}{lccccccc} \max\ z = & c_1 x_1 & +c_2 x_2 & +\cdots+ & c_n x_n & & & \\ & a_{11}x_1 & +a_{12}x_2 & +\cdots+ & a_{1n}x_n & \leq & b_1 & \leftarrow\ y_1 \\ & a_{21}x_1 & +a_{22}x_2 & +\cdots+ & a_{2n}x_n & \leq & b_2 & \leftarrow\ y_2 \\ & \vdots & & \ddots & & \vdots & & \\ & a_{m1}x_1 & +a_{m2}x_2 & +\cdots+ & a_{mn}x_n & \leq & b_m & \leftarrow\ y_m \\ & x_1, & x_2, & \cdots & x_n & \geq & 0 & \end{array}$$

How does a “small” change in b_i affect the total optimal value?

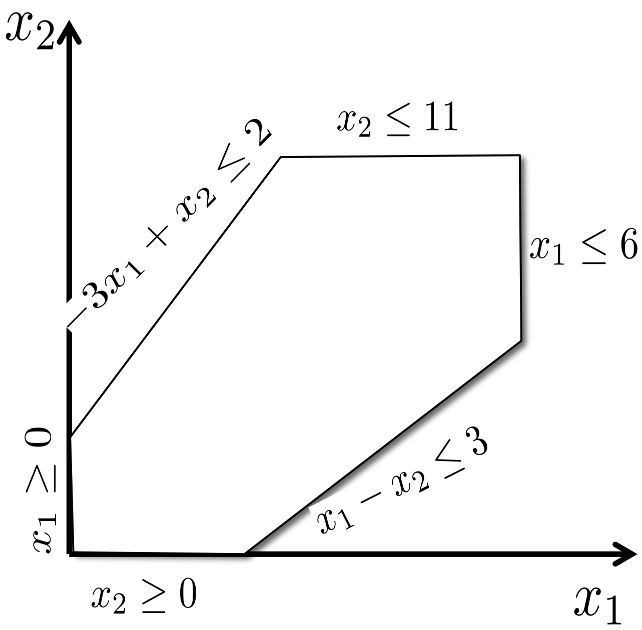
Linear Programming Problem

$$\begin{array}{rrrcl}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}$$



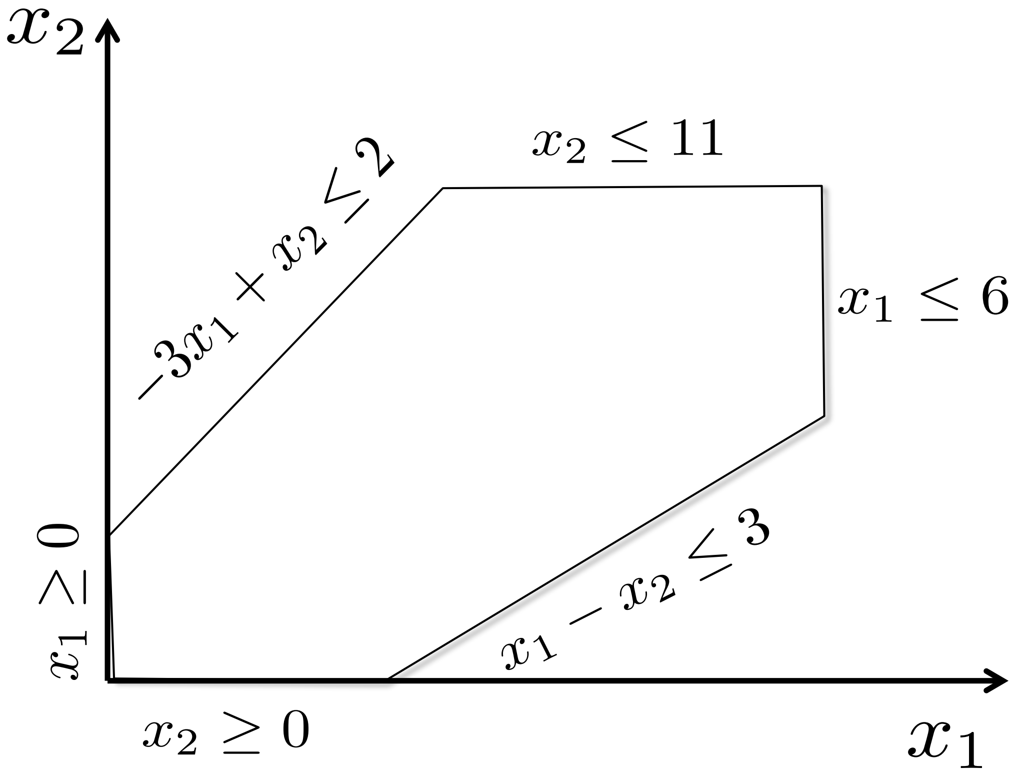
Dual Optimum

$$\begin{array}{lrrcrl}
\text{max.} & x_1 & +2x_2 & & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 & y_1 : 0 \\
 & & +x_2 & \leq & 11 & y_2 : 2 \\
 & x_1 & -x_2 & \leq & 3 & y_3 : 0 \\
 & x_1 & & \leq & 6 & y_4 : 1 \\
 & x_1, & x_2 & \geq & 0 & {\color{red} y_5, y_6} : 0
\end{array}$$



Geometric View

$$
\begin{array}{rrrcr}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}
\quad
\boxed{\begin{array}{rcl}
y_1 & : & 0 \\
y_2 & : & 2 \\
y_3 & : & 0 \\
y_4 & : & 1 \\
{\color{red}y_5, y_6} & : & 0
\end{array}}
$$



Sensitivity Analysis

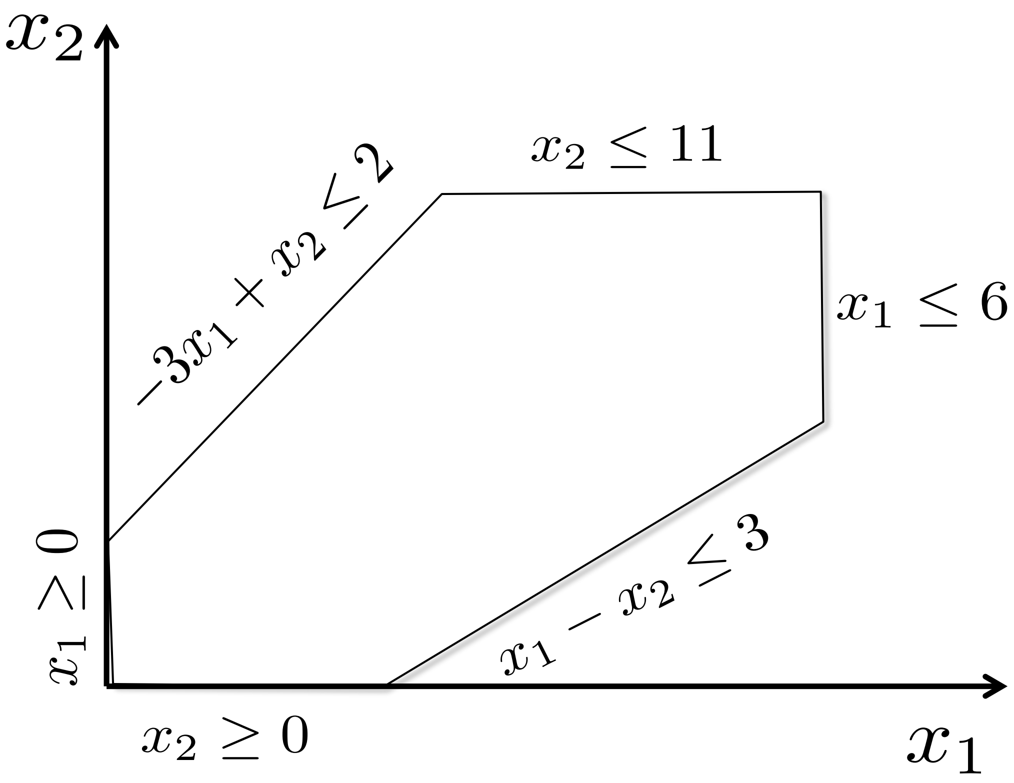
$$\begin{array}{rcccccccl} \max\ z = & c_1x_1 & +c_2x_2 & +\cdots+ & c_nx_n & & & & \\ & a_{11}x_1 & +a_{12}x_2 & +\cdots+ & a_{1n}x_n & \leq & b_1 & \leftarrow & y_1 \\ & a_{21}x_1 & +a_{22}x_2 & +\cdots+ & a_{2n}x_n & \leq & b_2 & \leftarrow & y_2 \\ & \vdots & & \ddots & & \vdots & & & \\ & a_{m1}x_1 & +a_{m2}x_2 & +\cdots+ & a_{mn}x_n & \leq & b_m & \leftarrow & y_m \\ & x_1, & x_2, & \cdots & x_n & \geq & 0 & & \end{array}$$

1. x^* and y^* be optimal primal/dual from final dictionary
2. dictionary is assumed non-degenerate.

For a *infinitesimally* small change d in b_j (I.e, b_j changes to b_j +d)

the objective changes by y_j * d

$$
\begin{array}{lrrcl}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}
\qquad
\boxed{\begin{array}{rcl}
y_1 & : & 0 \\
y_2 & : & 2 \\
y_3 & : & 0 \\
y_4 & : & 1 \\
\color{red}{y_5, y_6} & : & 0
\end{array}}
$$



Diet Problem Dual

$$\begin{array}{lrcl} \max & -\mathbf{c}^\intercal \mathbf{x} & & \\ & F^\intercal \mathbf{x} & \leq & \mathbf{u} \\ & -F^\intercal \mathbf{x} & \leq & -\ell \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

$$\begin{array}{lrcl} \min & \mathbf{u}^\intercal \mathbf{y}_\mathbf{u} - \ell^\intercal \mathbf{y}_\ell & & \\ & F(\mathbf{y}_\mathbf{u} - \mathbf{y}_\ell) & \geq & -\mathbf{c} \\ & \mathbf{y}_\mathbf{u},\ \mathbf{y}_\ell & \geq & \mathbf{0} \end{array}$$

DIET PROBLEM: DUAL

What does the dual mean?

Food	Calories	Total_Fat	Protein	Vit_A	Vit_C	Calcium	Price
Peppers	20	0.1	0.7	467.7	66.1	6.7	0.8
Potatoes, Baked	171.5	0.2	3.7	0	15.6	22.7	0.5
Tofu	88.2	5.5	9.4	98.6	0.1	121.8	1.1
Couscous	100.8	0.1	3.4	0	0	7.2	1
White Rice	102.7	0.2	2.1	0	0	7.9	0.4
Macaroni,Ckd	98.7	0.5	3.3	0	0	4.9	0.2
Peanut Butter	188.5	16	7.7	0	0	13.1	0.6

Nutrient	Min	Max
Calories	2000	2250
Total_Fat	0	65
Protein	50	100
Vit A	5000	50000
Vit C	50	20000
Calcium	800	1600

Optimal Solutions

Primal Solution

Food	Opt. Amt.
Peppers	9.55
Potatoes, Baked	0.95
Tofu	5.39
Couscous	0.00
White Rice	0.00
Macaroni,Ckd	11.86
Peanut Butter	0.00

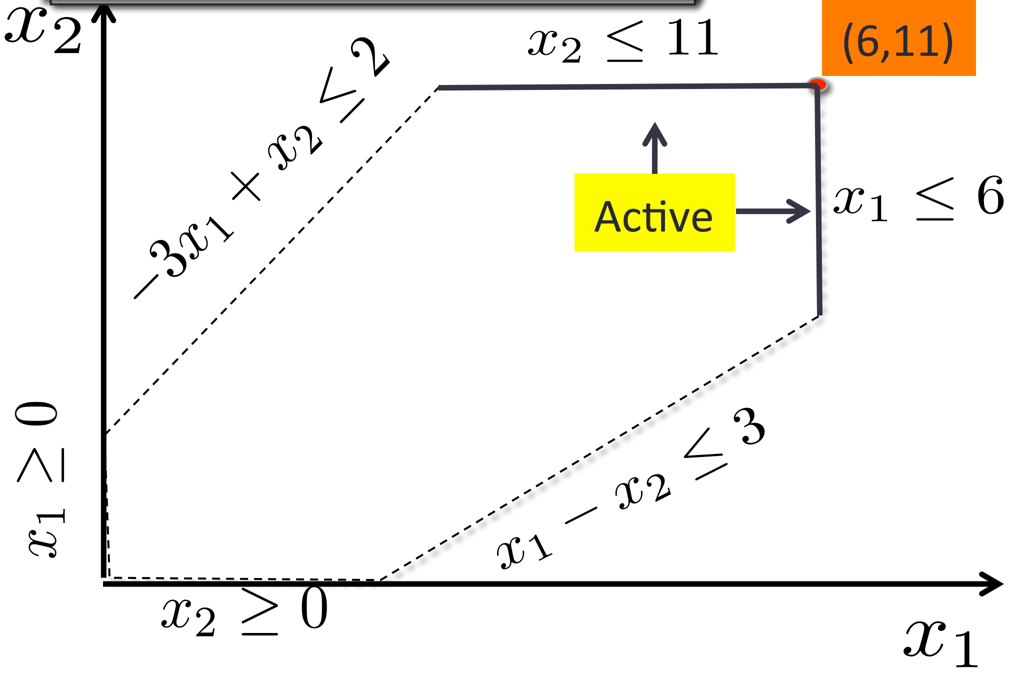
Dual Solution

Nutrient	Dual (yU)	Dual (yL)
Calories	0.000	0.002
Total_Fat	0.000	0.000
Protein	0.021	0.000
Vit A	0.000	0.002
Vit C	0.000	0.000
Calcium	0.000	0.008

COMPLEMENTARY SLACKNESS THEOREM

$$
\begin{aligned}
x_3 &= 2 + 3x_1 - x_2 \\
&= 2 + 18 - 11 \\
&= 9
\end{aligned}
$$

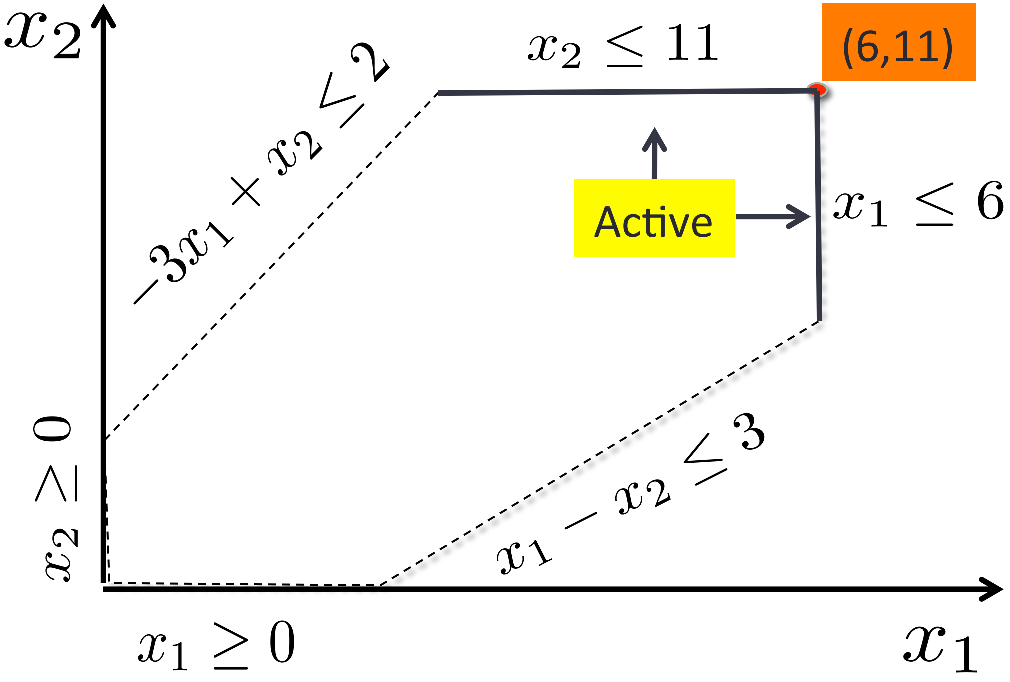
$$
\begin{array}{lrrcll}
\text{max.} & x_1 & +2x_2 & & & \\
x_3 \rightarrow & -3x_1 & +x_2 & \leq & 2 & \leftarrow y_1 \\
x_4 \rightarrow & & +x_2 & \leq & 11 & \leftarrow y_2 \\
x_5 \rightarrow & x_1 & -x_2 & \leq & 3 & \leftarrow y_3 \\
x_6 \rightarrow & x_1 & & \leq & 6 & \leftarrow y_4 \\
 & x_1, & x_2 & \geq & 0 & \leftarrow y_5, y_6
\end{array}
$$



Primal	Dual
$x_1 : 6$	$y_5 : 0$
$x_2 : 11$	$y_6 : 0$
$x_3 : 9$	$y_1 : 0$
$x_4 : 0$	$y_2 : 2$
$x_5 : 8$	$y_3 : 0$
$x_6 : 0$	$y_4 : 1$

Active vs. Inactive Constraints

Primal	Dual
$x_1 : 6$	$y_5 : 0$
$x_2 : 11$	$y_6 : 0$
$x_3 : 9$	$y_1 : 0$
$x_4 : 0$	$y_2 : 2$
$x_5 : 8$	$y_3 : 0$
$x_6 : 0$	$y_4 : 1$



$x_2 <= 11$ is active.

1. y_2 (dual) is non-zero.
2. x_4 (slack) is zero

$-3x_1 + x_2 <= 2$ is inactive.

1. y_1 (dual) is zero.
2. x_3 (slack) is non-zero

Complementary Slackness (Main Idea)

- Let **x** be a primal feasible solution
- Let **y** be a dual feasible solution.
- *Complementarity Condition:* Product of complementary pairs are all zero.

$$x_i \times y_j = 0$$

Complementary Pairs

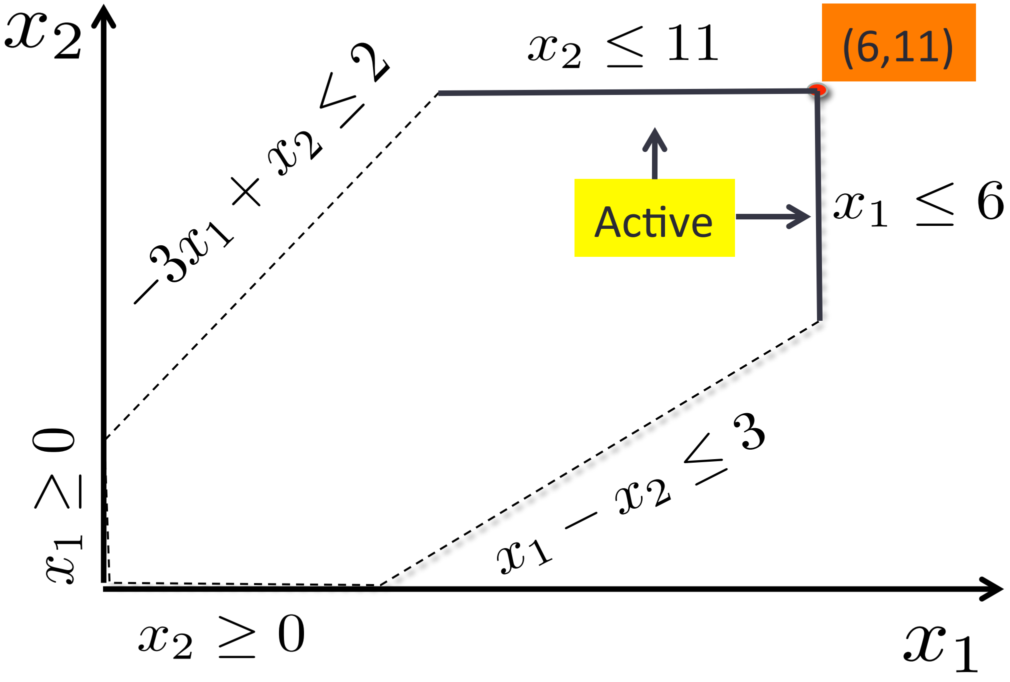
- Theorem: **x, y** are primal and dual optimal respectively.

Proof

Will be discussed separately on a discussion forum.

An Example

$$
\begin{array}{rrrrrl}
\text{max.} & & x_1 & +2x_2 & & \\
x_3 \rightarrow & & -3x_1 & +x_2 & \leq & 2 \leftarrow y_1 \\
x_4 \rightarrow & & & +x_2 & \leq & 11 \leftarrow y_2 \\
x_5 \rightarrow & & x_1 & -x_2 & \leq & 3 \leftarrow y_3 \\
x_6 \rightarrow & & x_1 & & \leq & 6 \leftarrow y_4 \\
& & x_1, & x_2 & \geq & 0 \leftarrow y_5, y_6
\end{array}
$$



Primal	Dual
x_1 : 6	y_5 : 0
x_2 : 11	y_6 : 0
x_3 : 9	y_1 : 0
x_4 : 0	y_2 : 2
x_5 : 8	y_3 : 0
x_6 : 0	y_4 : 1

Complementary Slackness

$$\begin{array}{c|cc} \mathbf{x_B} & \mathbf{b} & +A\mathbf{x_I} \\ \hline z & z_0 & +\mathbf{c}^\intercal \mathbf{x_I} \end{array} \qquad \begin{array}{c|cc} \mathbf{x_I}^c & -\mathbf{c} & -A^\intercal \mathbf{x_B}^c \\ \hline d & -z_0 & -\mathbf{b}^\intercal \mathbf{x_B}^c \end{array}$$

$$x_I = \mathbf{0}, x_B = \mathbf{b} \qquad x_B^c = \mathbf{0}, x_I^c = -\mathbf{c}$$

Claim: The solutions represented by primal and dual dictionaries are complementary pairs.

KARUSH-KUHN-TUCKER (KKT) CONDITIONS FOR LP

Karush-Kuhn-Tucker Conditions

- Very important for many optimization problems.

$$\begin{array}{llcl} \max & \mathbf{c}^\intercal \mathbf{x} & & \\ & A\,\mathbf{x} + \mathbf{x_s} & = & \mathbf{b} \\ & \mathbf{x}, \mathbf{x_s} & \geq & \mathbf{0} \end{array}$$

Primal

$$\begin{array}{llcl} \min & \mathbf{b}^\intercal \mathbf{y} & & \\ & A^\intercal\,\mathbf{y} - \mathbf{y_s} & = & \mathbf{c} \\ & \mathbf{y}, \mathbf{y_s} & \geq & \mathbf{0} \end{array}$$

Dual

Necessary and Sufficient Conditions for optimal solution

$$(\mathbf{x}, \mathbf{x_s}, \mathbf{y}, \mathbf{y_s})$$

KKT conditions for Linear Programs

The primal-dual solution $(\mathbf{x}, \mathbf{x_s}, \mathbf{y}, \mathbf{y_s})$ is optimal iff it satisfies the following conditions:

$$\begin{array}{rcl} A\,\mathbf{x} + \mathbf{x_s} & = & \mathbf{b} \\ \mathbf{x}, \mathbf{x_s} & \geq & \mathbf{0} \end{array}$$

(x,x_s) is primal feasible

$$\begin{array}{rcl} A^{\intercal}\,\mathbf{y} - \mathbf{y_s} & = & \mathbf{c} \\ \mathbf{y}, \mathbf{y_s} & \geq & \mathbf{0} \end{array}$$

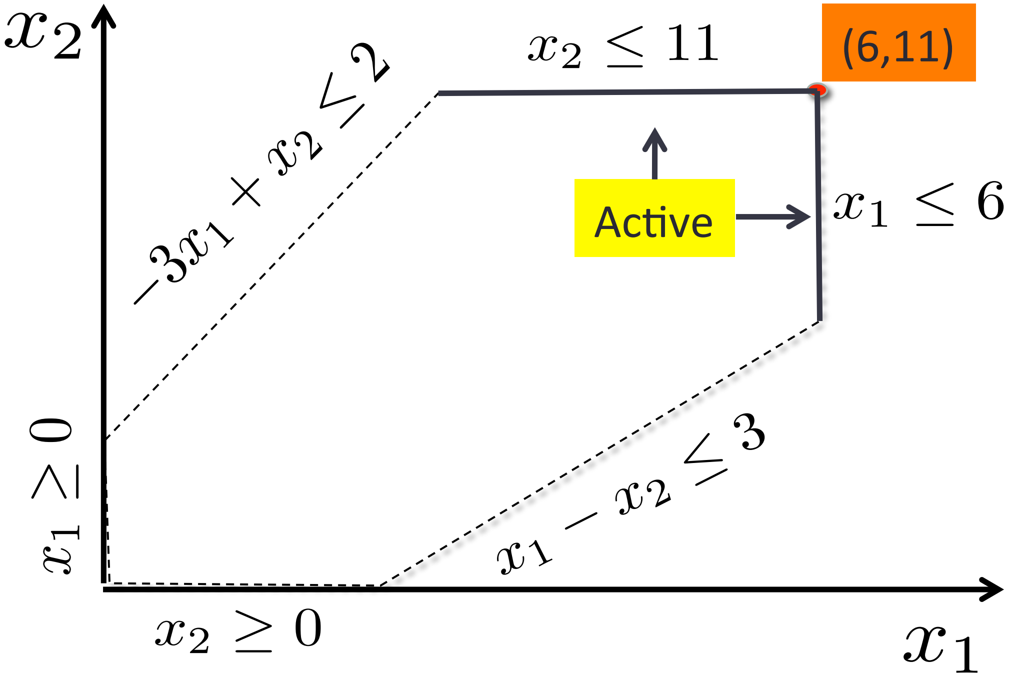
(y,y_s) is dual feasible

$$\begin{array}{l} x_j y_{s,j} = 0 \\ y_j x_{s,j} = 0 \end{array}$$

Product of all complementary pairs is zero.

An Example

$$\begin{array}{rrrcrl}
\text{max.} & x_1 & +2x_2 & & & \\
x_3 \rightarrow & -3x_1 & +x_2 & \leq & 2 & \leftarrow y_1 \\
x_4 \rightarrow & & +x_2 & \leq & 11 & \leftarrow y_2 \\
x_5 \rightarrow & x_1 & -x_2 & \leq & 3 & \leftarrow y_3 \\
x_6 \rightarrow & x_1 & & \leq & 6 & \leftarrow y_4 \\
& x_1, & x_2 & \geq & 0 & \leftarrow y_5, y_6
\end{array}$$



Primal	Dual
$x_1 : 6$	$y_5 : 0$
$x_2 : 11$	$y_6 : 0$
$x_3 : 9$	$y_1 : 0$
$x_4 : 0$	$y_2 : 2$
$x_5 : 8$	$y_3 : 0$
$x_6 : 0$	$y_4 : 1$

Final Dictionary and KKT conditions

$$\begin{array}{c|cc} \mathbf{x_B} & \mathbf{b} & +A\mathbf{x_I} \\ \hline z & z_0 & +\mathbf{c}^\intercal \mathbf{x_I} \end{array} \qquad \begin{array}{c|cc} \mathbf{x_I}^c & -\mathbf{c} & -A^\intercal \mathbf{x_B}^c \\ \hline d & -z_0 & -\mathbf{b}^\intercal \mathbf{x_B}^c \end{array}$$

$$x_I = \mathbf{0}, x_B = \mathbf{b} \qquad x_B^c = \mathbf{0}, x_I^c = -\mathbf{c}$$

Claim: The solutions represented by final primal and dual dictionaries satisfy the KKT condition!

DUAL FOR GENERAL FORM PROBLEMS

General Form LP

$$\begin{array}{llcl} \max & \mathbf{c}^\intercal \mathbf{x} & & \\ & P\mathbf{x} & = & \mathbf{q} \\ & A\mathbf{x} & \leq & \mathbf{b} \end{array}$$

$$\begin{array}{llclcl} \max & \mathbf{c}^\intercal \mathbf{x} & & & & \\ & P\mathbf{x} & \leq & \mathbf{q} & \leftarrow & \color{red}{\mathbf{r}} \\ & -P\mathbf{x} & \leq & -\mathbf{q} & \leftarrow & \color{red}{\mathbf{s}} \\ & A\mathbf{x} & \leq & \mathbf{b} & \leftarrow & \color{red}{\mathbf{y}} \end{array}$$

Dual Derivation

$$
\begin{array}{lccccc}
\max & \mathbf{c}^\intercal \mathbf{x} & & & & \\
 & P\mathbf{x} & \leq & \mathbf{q} & \leftarrow & \mathbf{r} \\
 & -P\mathbf{x} & \leq & -\mathbf{q} & \leftarrow & \mathbf{s} \\
 & A\mathbf{x} & \leq & \mathbf{b} & \leftarrow & \mathbf{y}
\end{array}
$$

Dual

$$
\begin{array}{lll}
\min & \mathbf{q}^\intercal\mathbf{r} - \mathbf{q}^\intercal\mathbf{s} + \mathbf{b}^\intercal\mathbf{y} & \\
\text{s.t.} & P^\intercal\mathbf{r} - P^\intercal\mathbf{s} + A^\intercal\mathbf{y} = \mathbf{c} & \\
 & \mathbf{r}, \mathbf{s}, \mathbf{y} & \geq \quad 0
\end{array}
$$

General Form Dual

$$\begin{array}{llcll} \text{max} & \mathbf{c}^\intercal \, \mathbf{x} & & & \\ & P\mathbf{x} & = & \mathbf{q} & \leftarrow \mathbf{w} \\ & A\mathbf{x} & \leq & \mathbf{b} & \leftarrow \mathbf{y} \end{array}$$

$$\begin{array}{llcl} \min & \mathbf{q}^\intercal \mathbf{w} + \mathbf{b}^\intercal \mathbf{y} & & \\ \text{s.t.} & P^\intercal \mathbf{w} + A^\intercal \mathbf{y} = \mathbf{c} & & \\ & \mathbf{y} & \geq & 0 \end{array}$$

Example

$$
\begin{array}{lrrrcrl}
\max & 2x_1 & -3x_2 & +x_3 & & & \\
\text{s.t.} & x_1 & -x_2 & & = & 5 & \leftarrow w_1 \\
 & x_1 & -2x_2 & +x_3 & = & 3 & \leftarrow w_2 \\
 & x_1 & & & \leq & 6 & \leftarrow y_1 \\
 & & & x_3 & \leq & 5 & \leftarrow y_2
\end{array}
$$

$$
\begin{array}{lrrrrcr}
\min & 5w_1 & +3w_2 & +6y_1 & +5y_2 & & \\
\text{s.t.} & w_1 & +w_2 & +y_1 & & = & 2 \\
 & -w_1 & -2w_2 & & & = & -3 \\
 & & w_2 & & +y_2 & = & 1 \\
 & & & y_1, & y_2 & \geq & 0
\end{array}
$$

Complementary Pairs

Dual	Primal
y_1	$(6 - x_1)$
y_2	$(5 - x_3)$

KKT conditions

$$
\begin{array}{lrcll}
\max & \mathbf{c}^\intercal \mathbf{x} & & & \\
 & P\mathbf{x} & = & \mathbf{q} & \leftarrow \mathbf{w} \\
 & A\mathbf{x} & \leq & \mathbf{b} & \leftarrow \mathbf{y}
\end{array}
$$

$$
\begin{array}{ll}
\min & \mathbf{q}^\intercal \mathbf{w} + \mathbf{b}^\intercal \mathbf{y} \\
\text{s.t.} & P^\intercal \mathbf{w} + A^\intercal \mathbf{y} = \mathbf{c} \\
 & \mathbf{y} \geq 0
\end{array}
$$

$$
\begin{array}{rcll}
P\mathbf{x} & = & \mathbf{q} & \\
A\mathbf{x} & \leq & \mathbf{b} & \texttt{Primal Feas.} \\
 & & & \\
P^\intercal \mathbf{w} + A^\intercal \mathbf{y} & = & \mathbf{c} & \\
\mathbf{y} & \geq & 0 & \texttt{Dual Feas.} \\
y_j(A_j\mathbf{x} - b_j) & = & 0 & j = 1, \ldots, m \\
 & & & \texttt{Complementarity}
\end{array}
$$

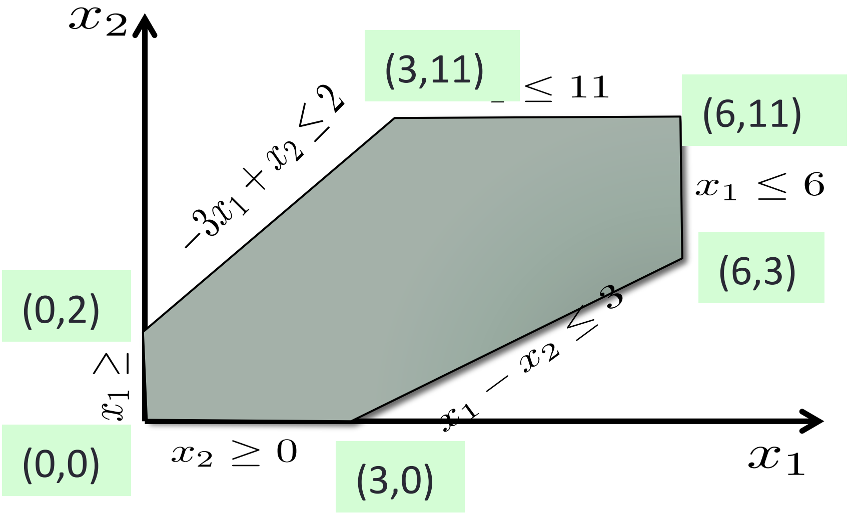
FINDING DUAL SOLUTION FROM DICTIONARY

Linear Programming Problem

From Two Weeks Ago.

$$
\begin{array}{llllll}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}
$$

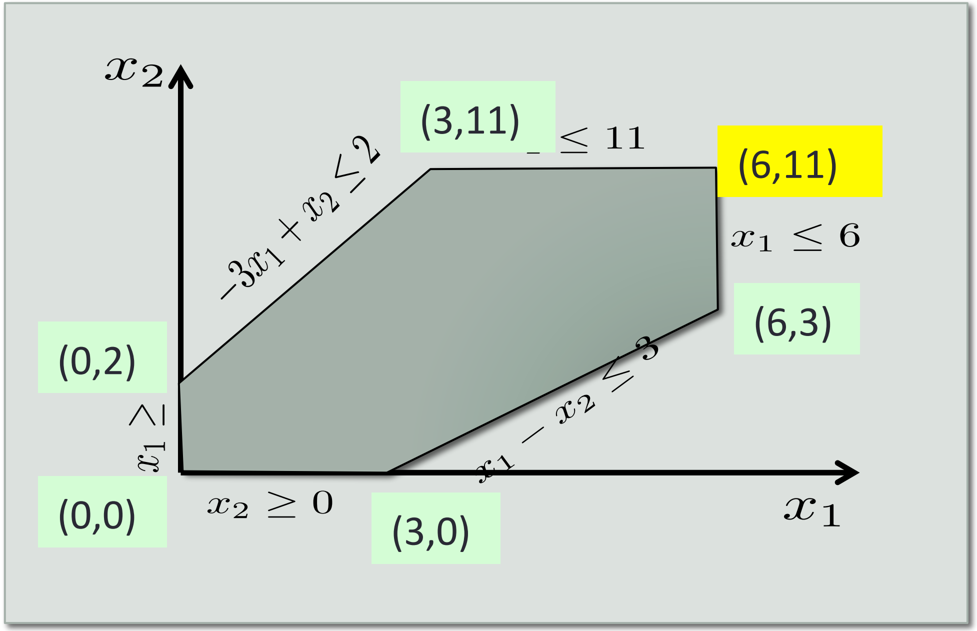
Note: Not drawn to scale



Goal: Solve LP using Simplex and visualize!

Final Dictionary

$$
\begin{array}{rcrrr}
x_3 & = & 9 & +x_4 & -3x_6 \\
x_1 & = & 6 & & -x_6 \\
x_2 & = & 11 & -x_4 & +0x_6 \\
x_5 & = & 8 & -x_4 & +x_6 \\
\hline
z & = & 28 & -2x_4 & -x_6
\end{array}
$$



How to read off the dual solution

$$\begin{array}{rcrrr} x_3 & = & 9 & +x_4 & -3x_6 \\ x_1 & = & 6 & & -x_6 \\ x_2 & = & 11 & -x_4 & +0x_6 \\ x_5 & = & 8 & -x_4 & +x_6 \\ \hline z & = & 28 & -2x_4 & -x_6 \end{array}$$

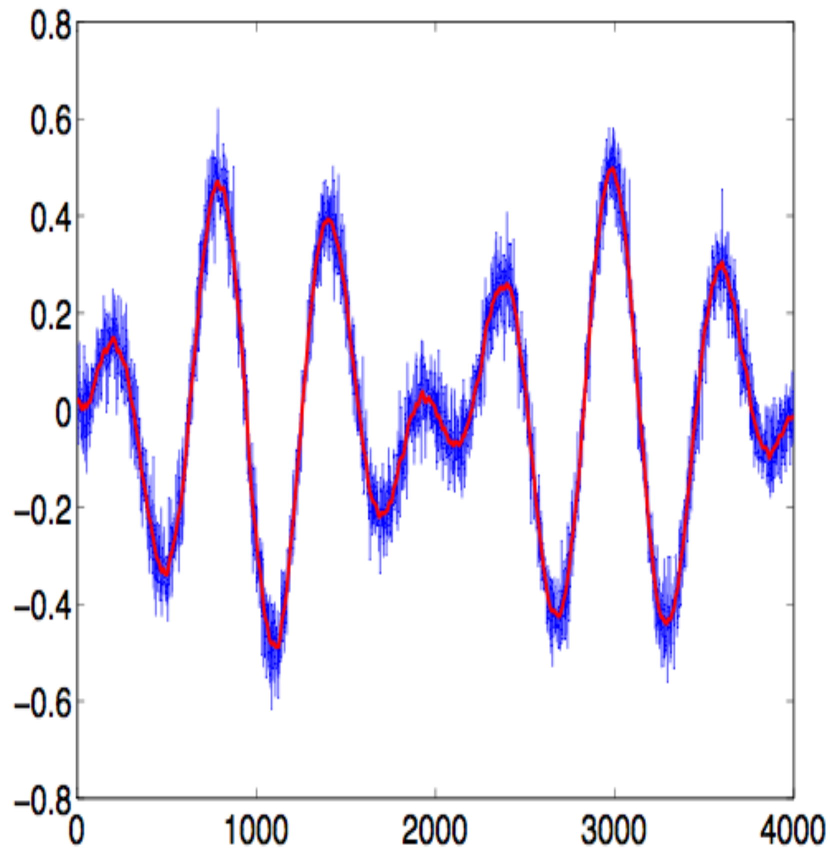
x_1	x_2	x_3	x_4	x_5	x_6
y_5	y_6	y_1	y_2	y_3	y_4

Reading off the dual from the final dictionary

$$\begin{array}{c|cc} \mathbf{x_B} & \mathbf{b} & +A\mathbf{x_I} \\ \hline z & z_0 & +\mathbf{c}^\intercal \mathbf{x_I} \end{array} \qquad \begin{array}{c|cc} \mathbf{x_I}^c & -\mathbf{c} & -A^\intercal \mathbf{x_B}^c \\ \hline d & -z_0 & -\mathbf{b}^\intercal \mathbf{x_B}^c \end{array}$$

Least-Squares

- Problem Description
- Overdetermined Problem
- "Best" Fit
- Matrix Form Solution
- De-Noising Application

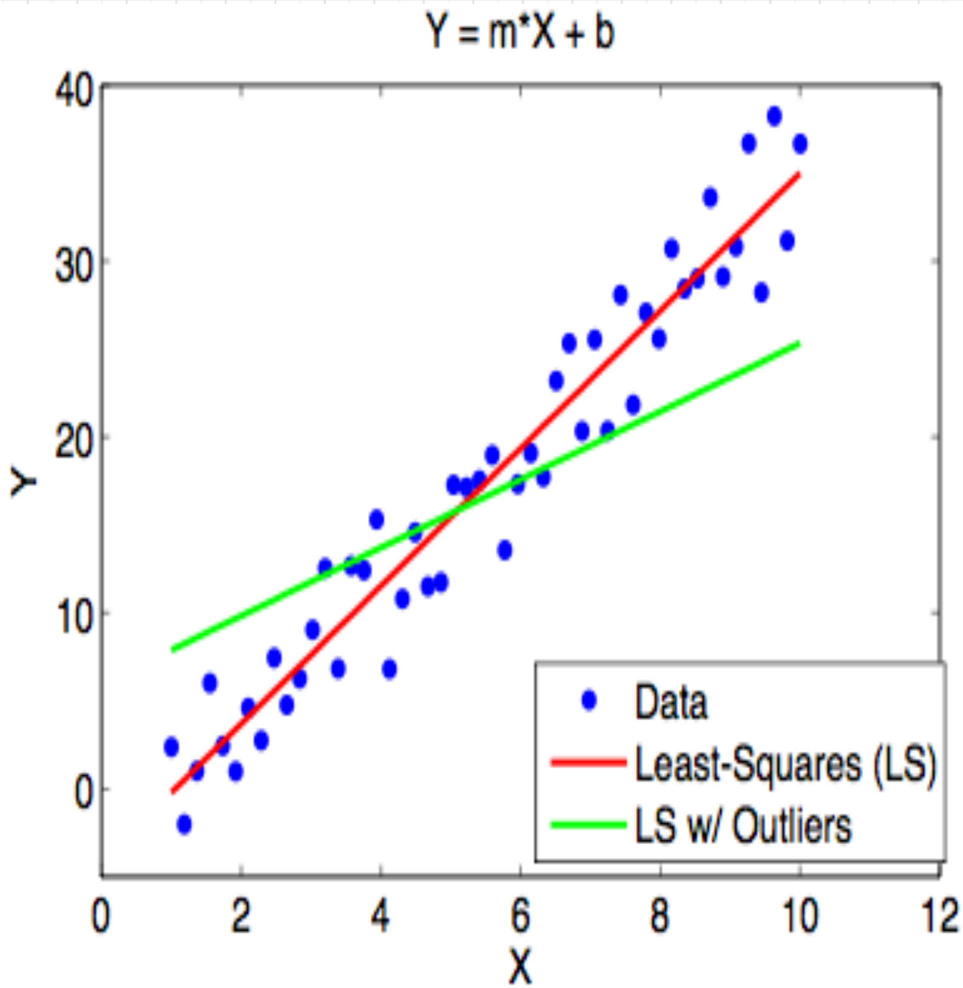


Problem Description

- n data points (x_i, y_i)

- Questions: Can you fit a line ($y = mx_i + b$) through all the data points?

In General, No!



Overdetermined Problem

- Too many equations, not enough unknowns

- No Solution! ✓
  (In General)

$mx_1 + b = y_1$

$mx_2 + b = y_2$

$\vdots$

$$\underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ b \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{b}$$

$$\boxed{A\boldsymbol{x} = \boldsymbol{b}}$$

Matrix Form Solution

- Matrix Solution: Solve the Normal Equation: $(A^T A)\boldsymbol{x}^\star = A^T \boldsymbol{b}$

$$\boxed{\boldsymbol{x}^\star = \underbrace{(A^T A)^{-1} A^T}_{+} \boldsymbol{b}}$$

A^+ - psuedo-inverse

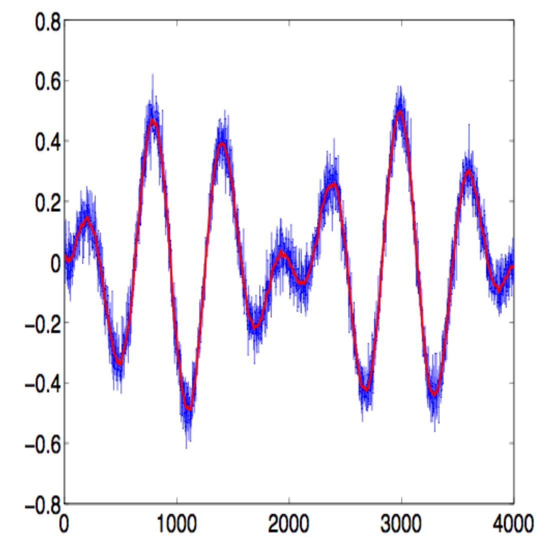
- In Matlab

```
>> x_star = pinv(A)*b;

>> x_star = A\b;
```

left divide

De-Noising Application



- Given n corrupted, by noise, data points

- Find n new data points that are:

    1.) Similar to the corrupted data

    2.) But smoother (i.e. the difference between neighboring data points should be small

1.)

2.)

$$\text{minimize} \quad \|\boldsymbol{x} - \boldsymbol{x}_{\text{cor}}\|^2 + \mu \sum_{k=1}^{n-1} (x_{k+1} - x_k)^2$$

if $\mu \to 0$

$x \to x_{cor}$

if $\mu \to \infty$

$x \to$ constant

De-Noising Application

- Re-cast problem as Least-Squares

$$\text{minimize} \quad \|A\boldsymbol{x} - \boldsymbol{b}\|^2$$

$$(\overbrace{\underbrace{mx_1 + b}_{}}^{Y(x_1)} - y_1)^2 + (\overbrace{mx_1 + b}^{Y(x_2)} - y_1)^2 + \cdots + (mx_n + b - y_n)^2$$

$$(\underbrace{\begin{bmatrix} x_1 & 1 \end{bmatrix}}_{a_1^T} \begin{bmatrix} m \\ b \end{bmatrix} - y_1)^2 + (\underbrace{\begin{bmatrix} x_2 & 1 \end{bmatrix}}_{a_2^T} \begin{bmatrix} m \\ b \end{bmatrix} - y_2)^2 \cdots$$

$$(\boldsymbol{a}_1^T x - \underset{Y_1}{b_1})^2 + (\boldsymbol{a}_1^T x - \underset{Y_2}{b_1})^2 \cdots + (\boldsymbol{a}_n^T x - b_n)^2$$

$$A = \begin{bmatrix} \boldsymbol{a_1}^T \\ \boldsymbol{a_2}^T \\ \vdots \\ \boldsymbol{a_n}^T \end{bmatrix}$$

$$\text{minimize} \quad \|\boldsymbol{x} - \boldsymbol{x}_{\text{cor}}\|^2$$

$$(x_1 - x_{cor}(1))^2 + (x_2 - x_{cor}(2))^2 + \cdots$$

$$\Big(\underbrace{[1\,0\,0 \cdots 0]}_{a_1^T}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \underbrace{x_{cor}(1)}_{b_1}\Big)^2 + \Big(\underbrace{[0\,1\,0\,0 \cdots 0]}_{a_2^T}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \underbrace{x_{cor}(2)}_{b_2}\Big)^2 + \cdots$$

$$A = \begin{bmatrix} I^{n \times n} \\ \\ ? \end{bmatrix} \qquad b = \begin{bmatrix} x_{cor} \\ \\ ? \end{bmatrix}$$

$$\text{minimize} \quad \mu \sum_{k=1}^{n-1} (x_{k+1} - x_k)^2$$

inside

$$\left(\sqrt{\mu}\,(x_2 - x_1)\right)^2 + \left(\sqrt{\mu}\,(x_3 - x_2)\right)^2 + \cdots \left(\sqrt{\mu}\,(x_n - x_{n-1})\right)^2$$

$$\Big(\underbrace{\sqrt{\mu}\,[-1\ 1\ 0\ 0 \cdots 0]}_{a_1^T} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \underset{b_1}{0}\Big)^2 + \Big(\underbrace{\sqrt{\mu}\,[0\ -1\ 1\ 0 \cdots 0]}_{a_2^T} x - \underset{b_2}{0}\Big)^2$$

$$\text{minimize} \quad \|\boldsymbol{x} - \boldsymbol{x}_{\text{cor}}\|^2 + \mu \sum_{k=1}^{n-1} (x_{k+1} - x_k)^2 = \|Ax - b\|^2$$

$$\underline{A} = \begin{bmatrix} I^{n \times n} \\ \sqrt{\mu}\, D^{n-1 \times n} \end{bmatrix} \qquad \underline{b} = \begin{bmatrix} x_{cor}^{\,n \times 1} \\ 0^{\,n-1 \times 1} \end{bmatrix}$$

$$D = \begin{array}{c} \overbrace{\hphantom{\begin{matrix} -1 & 1 & & & 0 \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & \ddots & \\ 0 & & & -1 & 1 \end{matrix}}}^{n} \\ \left.\begin{bmatrix} -1 & 1 & & & 0 \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & \ddots & \ddots \\ 0 & & & -1 & 1 \end{bmatrix}\right\} n-1 \end{array}$$

Least-Norm

- Problem Description

"fall" vs "fat"

- Underdetermined Problem

- Which solution is "best"?

- Matrix Form Solution

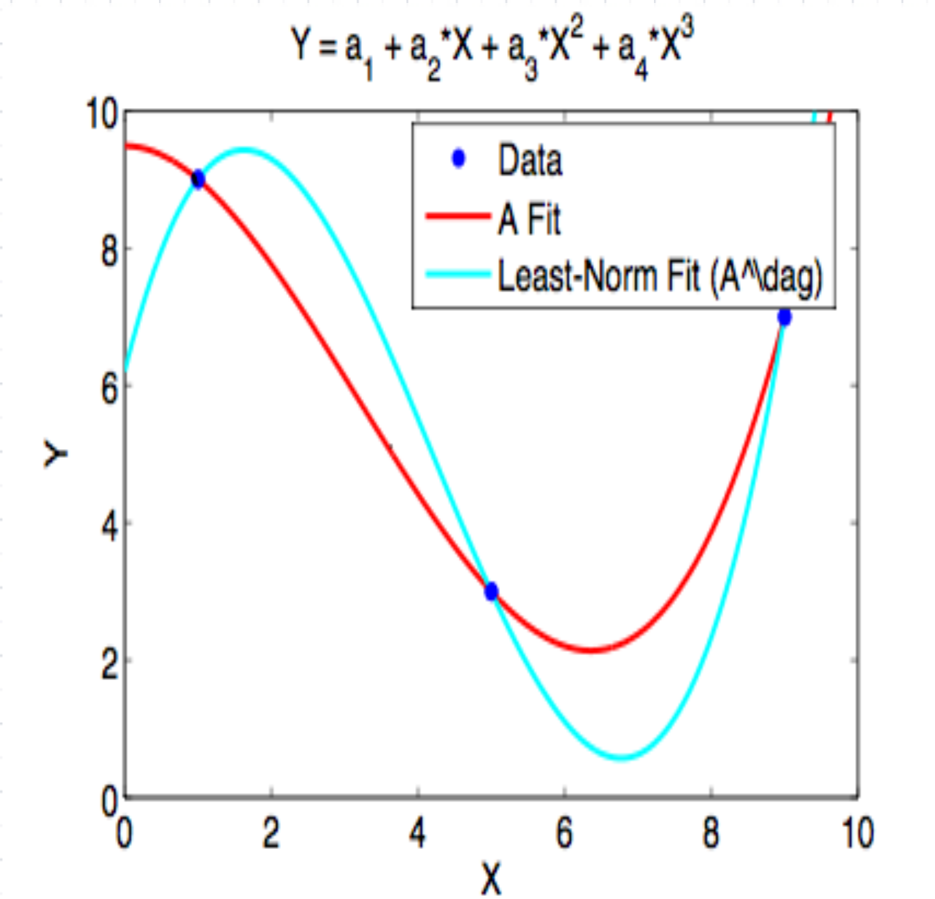
- Minimum Power Application

Problem Description

- m data points (x_i, y_i)

- Questions: Can you fit a polynomial ($y_i = a_0 + a_1 x_i + a_2 x_i^2 \cdots + a_n x_i^n$), with n coefficients, through all the data points?

$m < n$



Underdetermined Problem

- More than enough unknowns to solve the fewer equations ($n > m$)

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ & & \vdots & & \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

- Infinite Solutions!

$$Ax = b$$

$$a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 - y_i = 9$$

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Which Solution is "Best"

- All Residuals are zero

$$\boldsymbol{r} = A\boldsymbol{x} - \boldsymbol{b} = 0$$

- Choose the solutions with the "smallest" coefficients (i.e. smallest elements of the solution)

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{x}\|^2 = a_0^2 + a_1^2 + a_2^2 + a_3^2 \\ \text{subject to} & \boxed{A\boldsymbol{x} = \boldsymbol{b}} \end{array}$$

A closed-form Solution!

Matrix Form Solution

$\neq (A^T A)^{-1} A^T$ – least-squares

- Matrix Solution:

$$x^\star = \underbrace{A^T(AA^T)^{-1}}_{A^+}\boldsymbol{b}$$

A^+ – psuedo-inverse

- In Matlab

fast

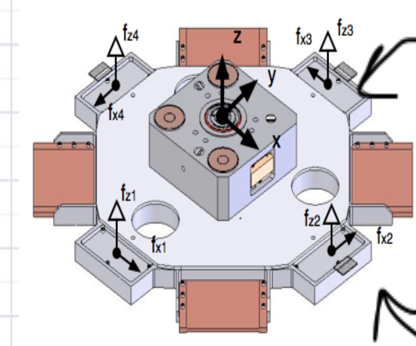
```
>> x_star = pinv(A)*b;
```

$x^* \neq A\backslash b$

x_star does not equal $\boxed{\texttt{A\textbackslash b}}$ (in general);

Minimum Power Application

Stage



motor

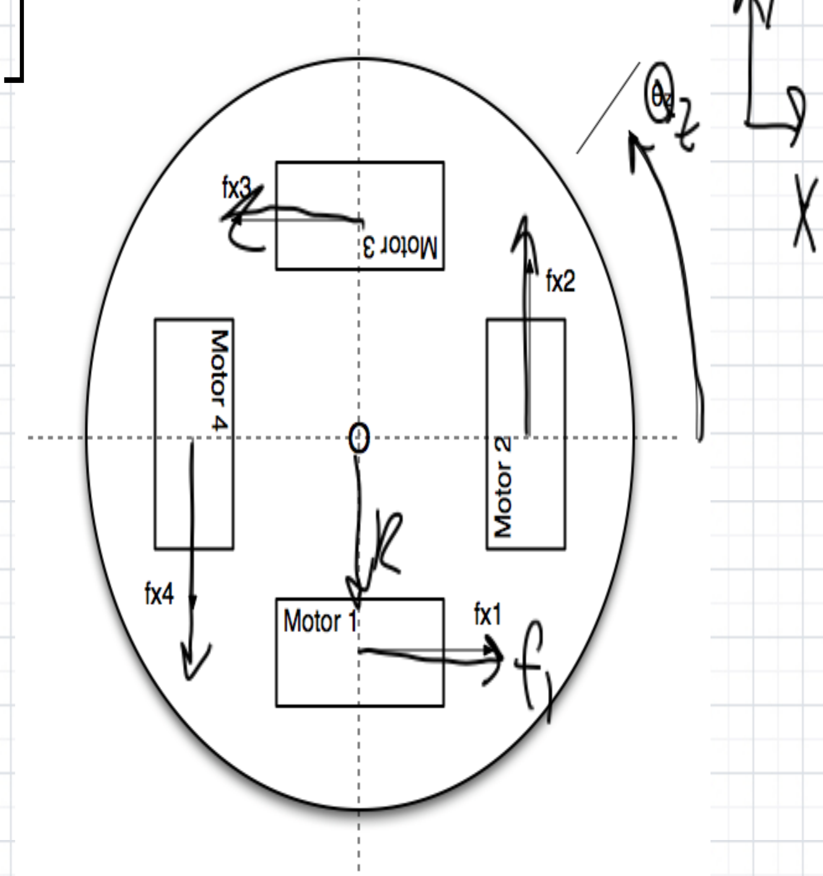
$$\begin{bmatrix} \boxed{F_x} \\ \boxed{F_y} \\ \boxed{\tau_z} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ R & R & R & R \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$$

Global forces

$$\boxed{\underset{3\times1}{\boldsymbol{F}} = A\underset{4\times1}{\boldsymbol{f}}}$$

$$F_x = f_1 - f_3$$

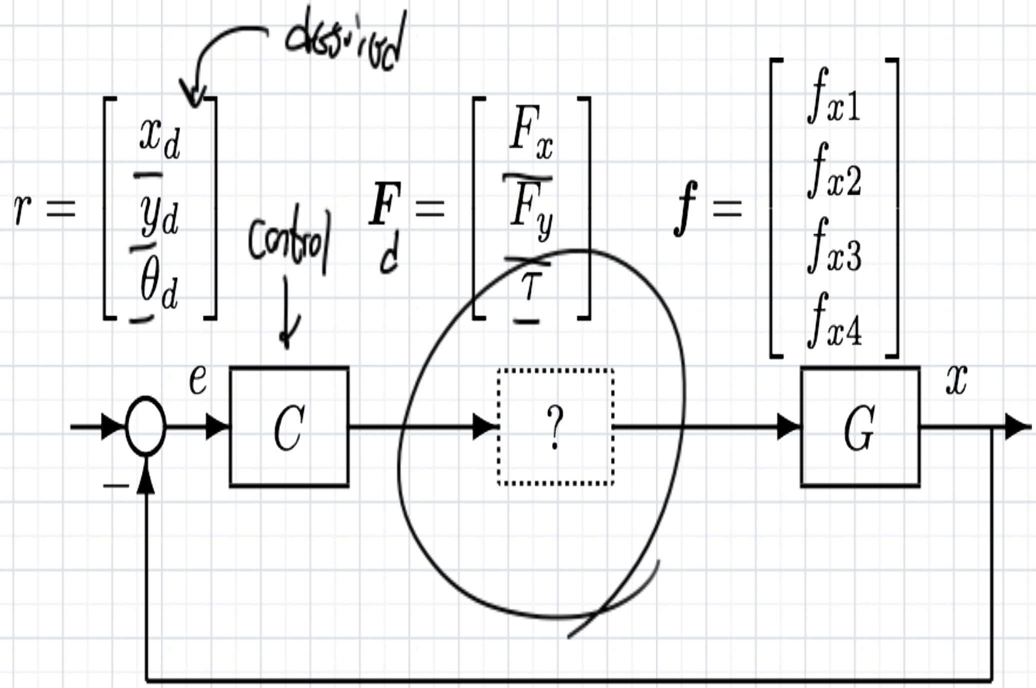
TOP VIEW



☆ Assuming θ is small

(small angle approximation)

Control Loop



3 eqns $\qquad$ 4 unknowns

$F \rightarrow f$

☆ Assume that force ≈ current (electromagnetic)

$\Rightarrow$ force^2 ≈ Power

current^2 ≈ Power

$$\min f_1^2 + f_2^2 + f_3^2 + f_4^2 = \|f\|_2^2 = \min \text{Power}$$

$$\text{s.t.} \quad \underline{F} = \underline{A} f$$

Least-Norm Solution

min $\|f\|_2^2$? min Power

s.t $F_d = Af$

Given

motor 1 $$f_1 = \frac{Fx}{2} + \frac{\tau_z}{4R}$$

$$f_2 = \frac{Fy}{2} + \frac{\tau_z}{4R}$$

motor 3 $$f_3 = -\frac{Fx}{2} + \frac{\tau_z}{4R}$$

$$f_4 = -\frac{Fy}{2} + \frac{\tau_z}{4R}$$

$$f^* = A^\dagger F_d = \begin{bmatrix} f_1^* \\ f_2^* \\ f_3^* \\ f_4^* \end{bmatrix} ?$$

Null Space of A

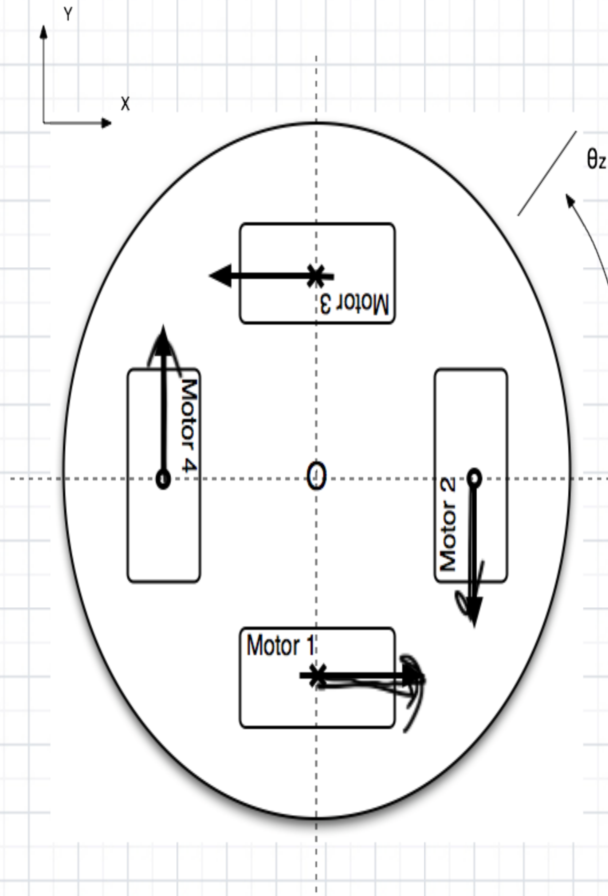
* Why are there an infinite set of solutions?

if $x_n \in N(A)$

$$\Rightarrow \quad A x_n = 0 = \text{zero!}$$

$$F = Af = A(\overbrace{f + f_N}^{\hat{f} \neq f}) \quad \text{where } f_N \in N(A)$$

$$= Af + \cancelto{0}{Af_N}$$



$f_1 = c$

$f_2 = -c$

$f_3 = c$

$f_4 = -c$

$$Af = 0$$

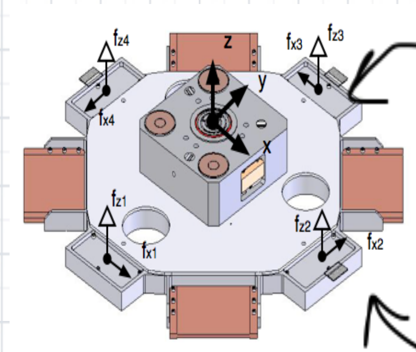
where $f \neq 0$

$$F_d = Af = A\Big(\underbrace{f + c\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}}_{\hat{f}}\Big) = Af$$

$$\hat{f} \neq f$$

Minimum Power Application

Stage



motor

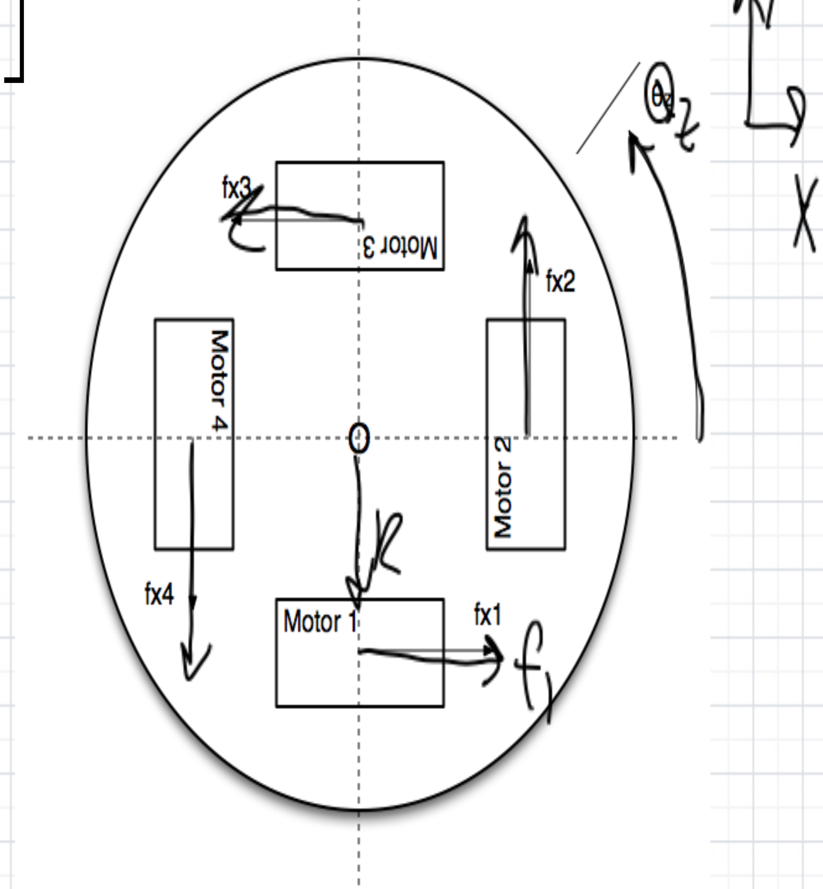
$$\begin{bmatrix} \boxed{F_x} \\ \boxed{F_y} \\ \boxed{\tau_z} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ R & R & R & R \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$$

$$\boxed{\boldsymbol{F}^{3x1} = A\boldsymbol{f}}^{\,4x1}$$

Global forces

$F_x = f_1 - f_3$

TOP VIEW



☆ Assuming θ is small

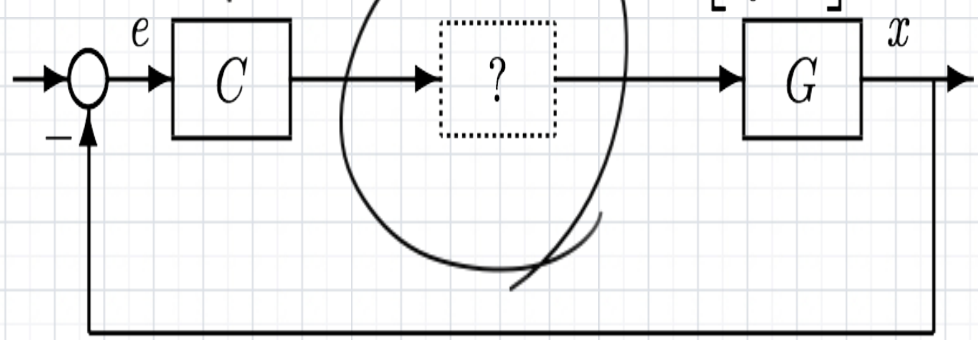
(small angle approximation)

Control Loop

desired

$$r = \begin{bmatrix} x_d \\ y_d \\ \theta_d \end{bmatrix} \qquad \boldsymbol{F}_d = \begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} \qquad \boldsymbol{f} = \begin{bmatrix} f_{x1} \\ f_{x2} \\ f_{x3} \\ f_{x4} \end{bmatrix}$$

control



3 Eqns $\quad$ 4 unknowns

$$F \longrightarrow f$$

★ Assume that force $\approx$ current (electromagnetic)

$$\text{current}^2 \approx \text{Power} \quad \Rightarrow \quad \text{force}^2 \approx \text{Power}$$

$$\min \; f_1^2 + f_2^2 + f_3^2 + f_4^2 = \|f\|_2^2 = \min \text{Power}$$

$$\text{s.t.} \quad F = \underline{A} f$$

Least-Norm Solution

min $\|f\|_2^2$ ⇒ min Power

s.t $F_d = Af$

Given

motor 1 $$f_1 = \frac{Fx}{2} + \frac{\tau_z}{4R}$$

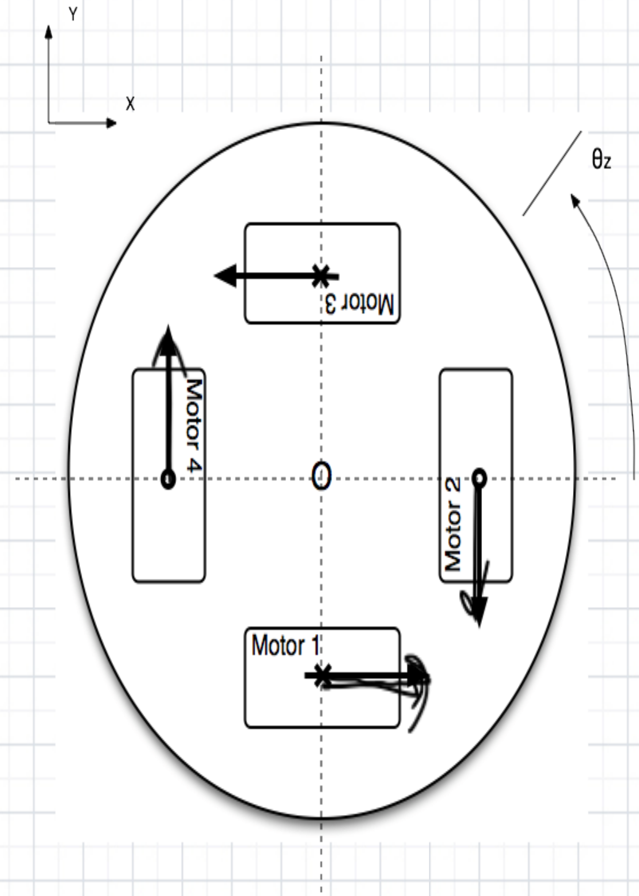
$$f_2 = \frac{Fy}{2} + \frac{\tau_z}{4R}$$

motor 3 $$f_3 = -\frac{Fx}{2} + \frac{\tau_z}{4R}$$

$$f_4 = -\frac{Fy}{2} + \frac{\tau_z}{4R}$$

$$f^* = A^\dagger F_d = \begin{bmatrix} f_1^* \\ f_2^* \\ f_3^* \\ f_3^* \end{bmatrix}$$

Null Space of A



$\ast$ Why are there an infinite set of Solutions!

if $\quad x_n \in N(A)$

$$\Rightarrow \quad A x_n = 0 = \text{zero!}$$

$$F = Af = A(\underbrace{f + f_N}_{\hat{f} \neq f}) \quad \text{where } f_N \in N(A)$$

$$= Af + \cancel{A f_N} \rightarrow 0$$

$$Af = 0$$

where $f \neq 0$

$f_1 = c$

$f_2 = -c$

$f_3 = c$

$f_4 = -c$

$$F_d = Af = A\left(\underbrace{f + c\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}}_{\hat{f} \neq f}\right) = Af$$

Piecewise - Linear Optimization

- Linear vs. Affine

- Piecewise - Linear (or Affine) Functions

- LP Equivalent (General)

- LP Equivalent $\underline{\ell_\infty}$ - Norm Minimization

- LP Equivalent $\underline{\ell_1}$ - Norm Minimization

Linear vs. Affine

- Linear:

$$f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + 0$$

- Affine:

$$f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + b$$

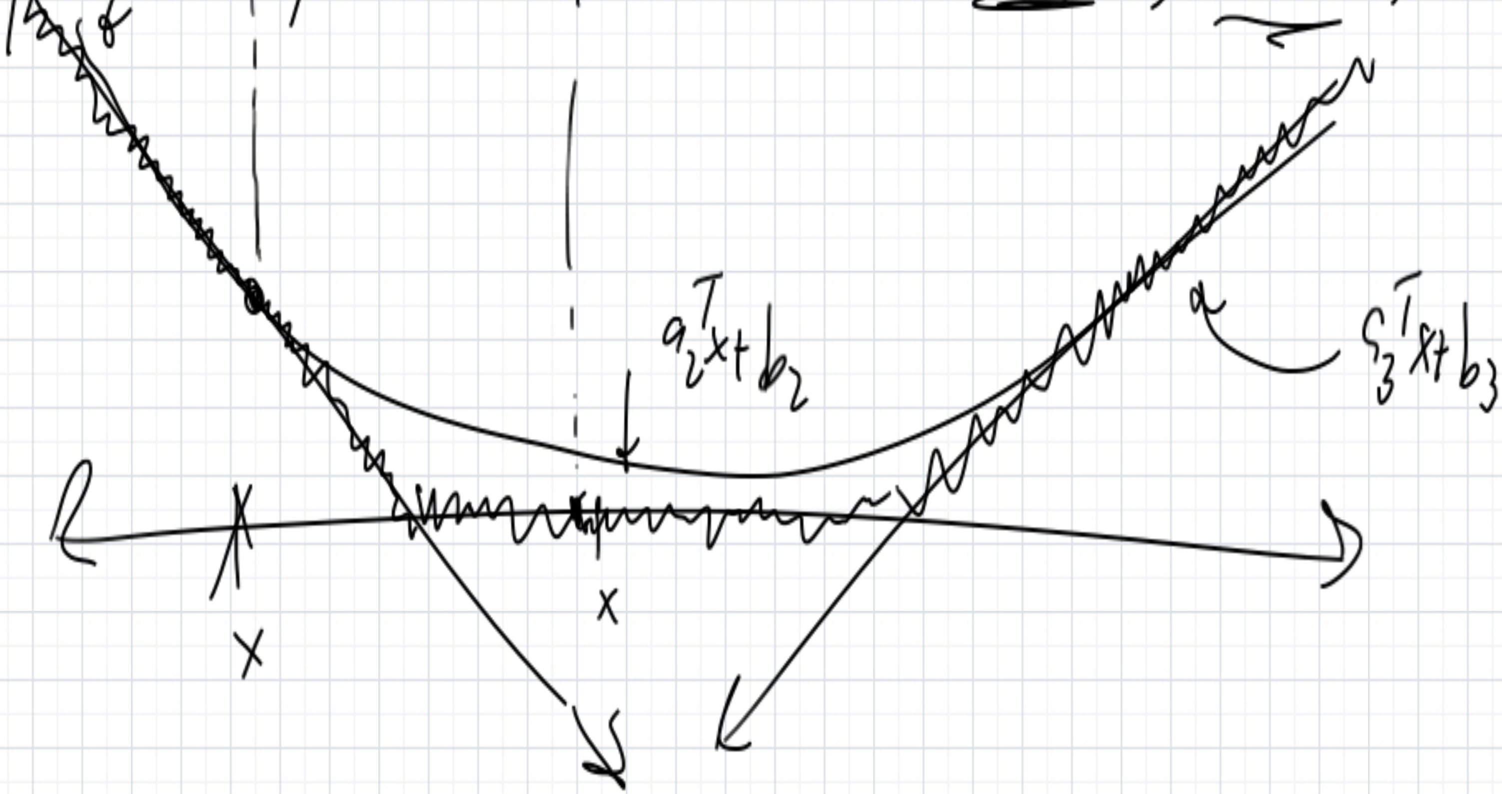
Piecewise - Linear (or Affine) Function

- Convex Function!

- Non - Differentiable Function.

$$f(\boldsymbol{x}) = \max_{i=1,2,\cdots,n} (\boldsymbol{a_i}^T \boldsymbol{x} + b_i)$$

$$= \max\{a_1^T x + b_1,\ a_2^T x + b_2,\ \cdots\}$$



LP Equivalent

$\min \ \max \{ a_1^T x + b_1, \ a_2^T x + b_2, \ \cdots \}$

- Need to add an auxiliary variable (scalar) t

$=$

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & \boldsymbol{a_i}^T \boldsymbol{x} + b_i \le t \end{aligned}$$

$a_1^T x + b_1 \le t$

$a_2^T x + b_2 \le t$

$\vdots$

$\hat{A} \hat{x} \le \hat{b}$

- Matrix Form:

$\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix}$ - new state-vector

$\min \ \underbrace{\begin{bmatrix} \vec{0} & 1 \end{bmatrix}}_{\hat{c}^T} \begin{bmatrix} x \\ t \end{bmatrix} = t$

s.t. $\begin{bmatrix} a_1^T & -1 \\ a_2^T & -1 \\ \vdots & -1 \end{bmatrix} \le \begin{bmatrix} -b_1 \\ -b_2 \\ \vdots \end{bmatrix}$

$$\min \underbrace{\begin{bmatrix} \vec{0} & 1 \end{bmatrix}}_{\hat{c}^T} \underbrace{\begin{bmatrix} x \\ t \end{bmatrix}}_{\hat{x}}$$

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix}$$

$$\text{s.t.} \quad \underbrace{\begin{bmatrix} A & -\mathbb{1} \end{bmatrix}}_{\hat{A}} \underbrace{\begin{bmatrix} x \\ t \end{bmatrix}}_{\hat{x}} \leq \underbrace{\begin{bmatrix} -b \end{bmatrix}}_{\hat{b}}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

>> linprog$(\hat{c}, \hat{A}, \hat{b})$

Sum of Piecewise- Linear Functions

Min $\underbrace{\max_{i=1,2,\cdots,n} (\boldsymbol{a_i}^T \boldsymbol{x} + b_i)}_{f(x)} + \underbrace{\max_{i=1,2,\cdots,m} (\boldsymbol{c_i}^T \boldsymbol{x} + d_i)}_{g(x)}$

- Equivalent LP
  - Need to add 2 auxiliary variables t_1, t_2

$$\begin{array}{ll} \text{minimize} & t_1 + t_2 \\ \text{subject to} & \boldsymbol{a_i}^T \boldsymbol{x} + b_i \leq t_1 \\ & \boldsymbol{c_i}^T \boldsymbol{x} + d_i \leq t_2 \end{array}$$

$$\left\{ \begin{array}{l} a_1^T x + b_1 \leq t_1 \\ \vdots \\ a_n^T x + b_n \leq t_1 \\ c_1^T x + d_1 \leq t_2 \\ \vdots \\ c_m^T x + d_m \leq t_2 \end{array} \right.$$

Sum of Piecewise- Linear Functions

- Matrix Form:

$$
\begin{aligned}
\text{minimize} \quad & t_1 + t_2 \\
\text{subject to} \quad & \boldsymbol{a_i}^T \boldsymbol{x} + b_i \le t_1 \\
& \boldsymbol{c_i}^T \boldsymbol{x} + d_i \le t_2
\end{aligned}
$$

$$\hat{A}\hat{x} \le \hat{b}$$

$$\hat{x} = \begin{bmatrix} x \\ t_1 \\ t_2 \end{bmatrix}$$

$$\min \underbrace{\begin{bmatrix} \vec{0} & 1 & 1 \end{bmatrix}^T}_{\hat{c}} \begin{bmatrix} x \\ t_1 \\ t_2 \end{bmatrix}$$

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} ; \quad C = \begin{bmatrix} c_1^T \\ \vdots \\ c_m^T \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} ; \quad d = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} A & -\mathbb{1} & \vec{0} \\ C & \vec{0} & -\mathbb{1} \end{bmatrix}}_{\hat{A}} \underbrace{\begin{bmatrix} x \\ t_1 \\ t_2 \end{bmatrix}}_{\hat{x}} \le \underbrace{\begin{bmatrix} -b \\ -d \end{bmatrix}}_{\hat{b}}$$

LP Equivalent: ℓ_∞ - Norm Approximation

- ℓ_∞ - Norm: $\quad ||\boldsymbol{y}||_\infty = \max_i |y_i| = \max_i \max(-y_i, y_i)$

  $|y_i|$

- Fitting/Approximation Problem: $\quad ||r||_\infty$

$$\text{minimize} \quad ||A\boldsymbol{x} - \boldsymbol{b}||_\infty$$

- LP Equivalent:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & A\boldsymbol{x} - \boldsymbol{b} \leq t\mathbf{1} \\ & -(A\boldsymbol{x} - \boldsymbol{b}) \leq t\mathbf{1} \end{aligned}$$

LP Equivalent: ℓ_∞ - Norm Approximation

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax - b \leq t\mathbf{1} \\ & -(Ax - b) \geq t\mathbf{1} \end{array}$$

$t\mathbf{1} = \begin{bmatrix} t \\ t \\ t \\ \vdots \end{bmatrix}$

$\min \|Ax - b\|_\infty$

$-Ax + b$

$\leq$

$\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix}$

$\min \underbrace{\begin{bmatrix} \vec{0} & 1 \end{bmatrix}^T}_{\hat{c}^T} \overbrace{\begin{bmatrix} x \\ t \end{bmatrix}}^{\hat{x}}$

$Ax - \mathbb{1}t \leq b$

$-Ax - \mathbb{1}t \leq -b$

ngisfalce

$-Ax + b \leq t\mathbb{1}$

$$\overbrace{\begin{bmatrix} A & -\mathbb{1} \\ A & -\mathbb{1} \end{bmatrix}}^{\hat{A}} \begin{bmatrix} x \\ t \end{bmatrix} \leq \overbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}^{\hat{b}}$$

LP Equivalent: ℓ_∞ - Norm Approximation

- Matrix Form:

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} \mathbf{0} & 1 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{x} \\ t \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \end{bmatrix} \leq \begin{bmatrix} \boldsymbol{b} \\ -\boldsymbol{b} \end{bmatrix} \end{array}$$

LP Equivalent: ℓ_1 - Norm Approximation

- ℓ_1 - Norm: $\quad ||\boldsymbol{y}||_1 = \sum_i |y_i| = \sum_i \max(-y_i, y_i)$ = $\underbrace{\max(-y_1, y_1)}_{f_1(x)\ \to\ t_1}$ $+ \underbrace{\max(-y_2, y_2)}_{f_2(x)\ \to\ t_2} \ldots$

- Fitting/Approximation Problem:

$$\text{minimize} \quad ||A\boldsymbol{x} - \boldsymbol{b}||_1$$

$A \in \mathbb{R}^{m \times n}$

- LP Equivalent:
  -- Need to add an auxiliary vector: $\quad \boldsymbol{t} = [\ t_1 \quad t_2 \quad \cdots \quad t_{m} \ ]^T$

m - # of Constraint equations

$$\begin{array}{ll} \text{minimize} & t_1 + t_2 + \cdots t_n \\ \text{subject to} & A\boldsymbol{x} - \boldsymbol{b} \ \leq \boldsymbol{t} \\ & -(A\boldsymbol{x} - \boldsymbol{b}) \geq \boldsymbol{t} \end{array}$$

$-Ax + b \leq t$ $\quad \leq$

$a_1^T x - b_1 \leq t_1$

$a_2^T x - b_1 \leq t_2$

$\vdots$

LP Equivalent: ℓ_1 - Norm Approximation

$$
\begin{array}{ll}
\text{minimize} & t_1 + t_2 + \cdots t_n = \mathbb{1}^T t \\
\text{subject to} & A\boldsymbol{x} - \boldsymbol{b} \le \boldsymbol{t} \\
& -(A\boldsymbol{x} - \boldsymbol{b}) \cancel{\geq} \le \boldsymbol{t}
\end{array}
$$

$\rightarrow Ax - It \le b$

$\rightarrow -Ax - It \le -b$

$$\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix} \Rightarrow \underbrace{\begin{bmatrix} \vec{0} & \mathbb{1} \end{bmatrix}}_{\hat{c}^T} \begin{bmatrix} x \\ t \end{bmatrix} = \mathbb{1}^T t = t_1 + t_2 + \cdots$$

$$\overbrace{\begin{bmatrix} A & -I \\ -A & -I \end{bmatrix}}^{\hat{A}} \overbrace{\begin{bmatrix} x \\ t \end{bmatrix}}^{\hat{x}} \le \overbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}^{\hat{b}}$$

$$\boxed{\begin{array}{ll} \min & \hat{c}^T \hat{x} \\ \text{s.t.} & \hat{A}\hat{x} \le \hat{b} \end{array}}$$

LP Equivalent: ℓ_1 - Norm Approximation

- Matrix Form:

$$
\begin{array}{ll}
\text{minimize} & \begin{bmatrix} \mathbf{0} & \mathbf{1} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{t} \end{bmatrix} \\
\text{subject to} & \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{t} \end{bmatrix} \leq \begin{bmatrix} \boldsymbol{b} \\ -\boldsymbol{b} \end{bmatrix}
\end{array}
$$

Piecewise - Linear Optimization

- Compare ℓ_2 vs ℓ_1 vs ℓ_∞ Minimization

    -- Histogram of Residuals (intuition)

    -- Line - Fitting Application

Solutions:

$\min \|\underbrace{Ax - b}_{r}\|_p$ where $p = \{1, 2, \infty\}$

- ℓ_2

$$\underline{x^\star = A^\dagger b}$$

– closed-form!

- ℓ_∞

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} \mathbf{0} & 1 \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} = t \\ \text{subject to} & \begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{array}$$

$$\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix} \begin{matrix} \}80 \\ \}+1 \end{matrix} \quad \underbrace{}_{81}$$

- ℓ_1

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} \mathbf{0} & \mathbf{1} \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{array}$$

$$\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix} \begin{matrix} \}80 \\ \}+200 \end{matrix} \quad \underbrace{}_{280}$$

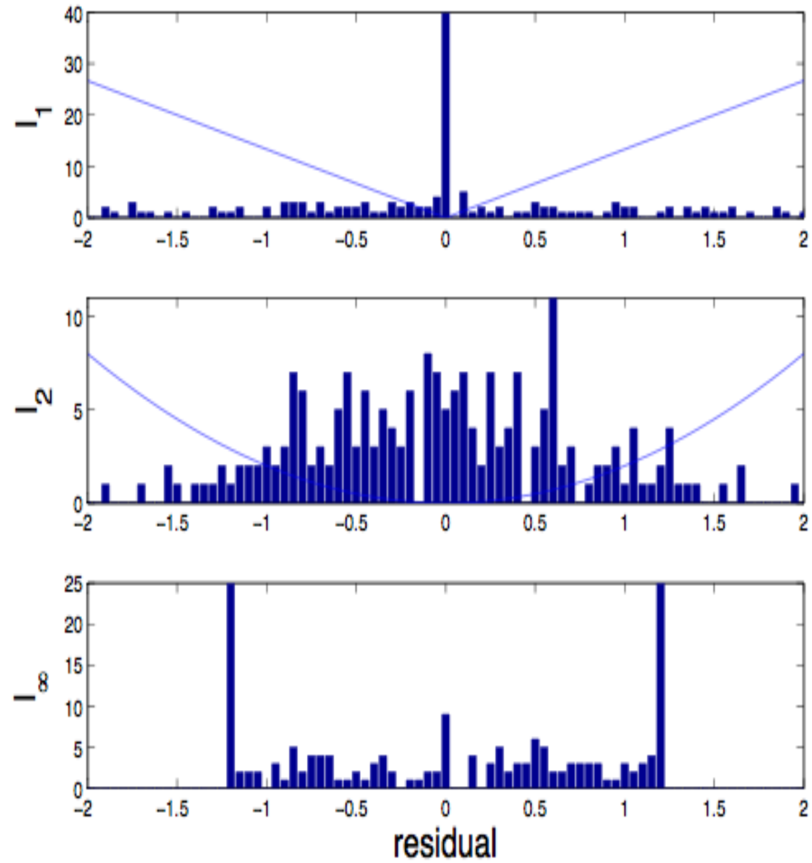
Approximation/Fitting Problem

- minimize $\|Ax - b\|_p$

- $A \in \mathbf{R}^{200\times 80}$ ← variables

eqns (constraints)

- Generate the elements of A and b via `randn( )` function



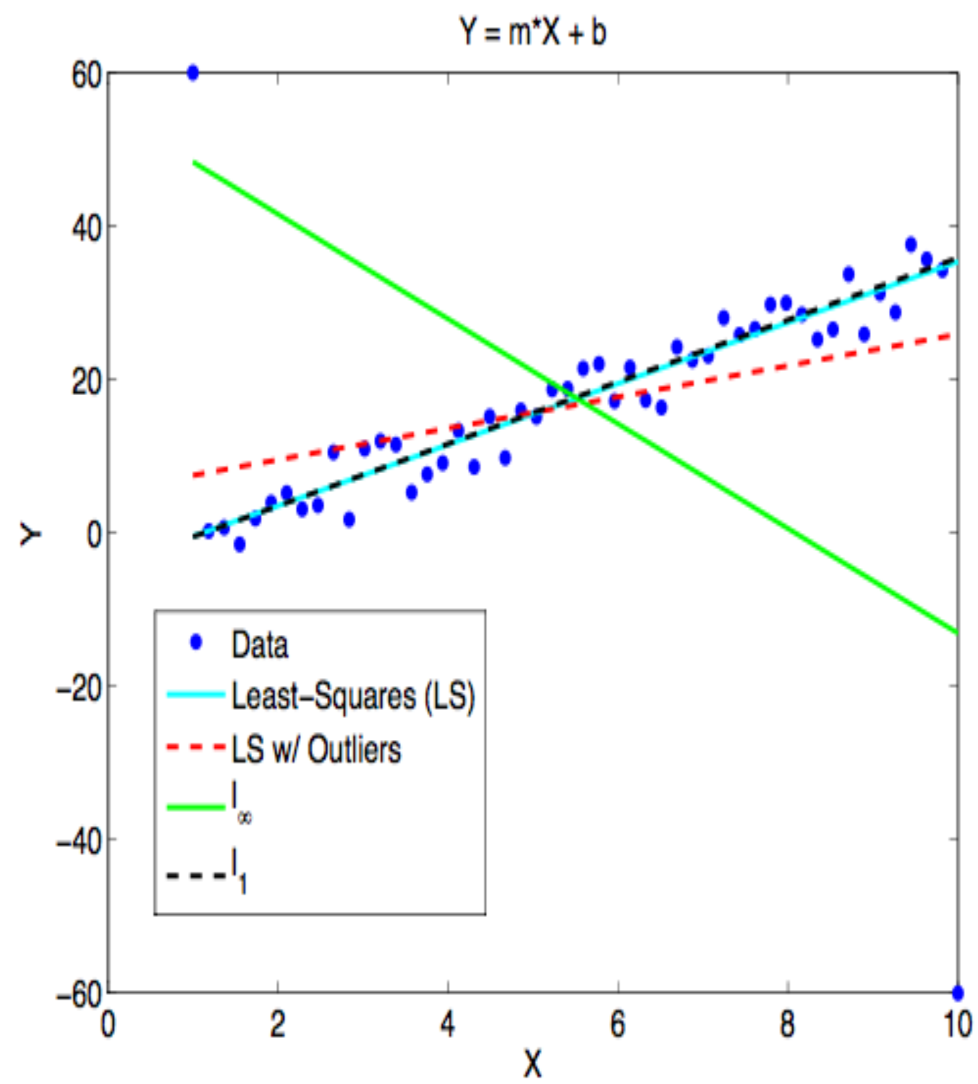
* Want!, then residuals = 0 !

Line Fitting Application

- Fit a line through data points.

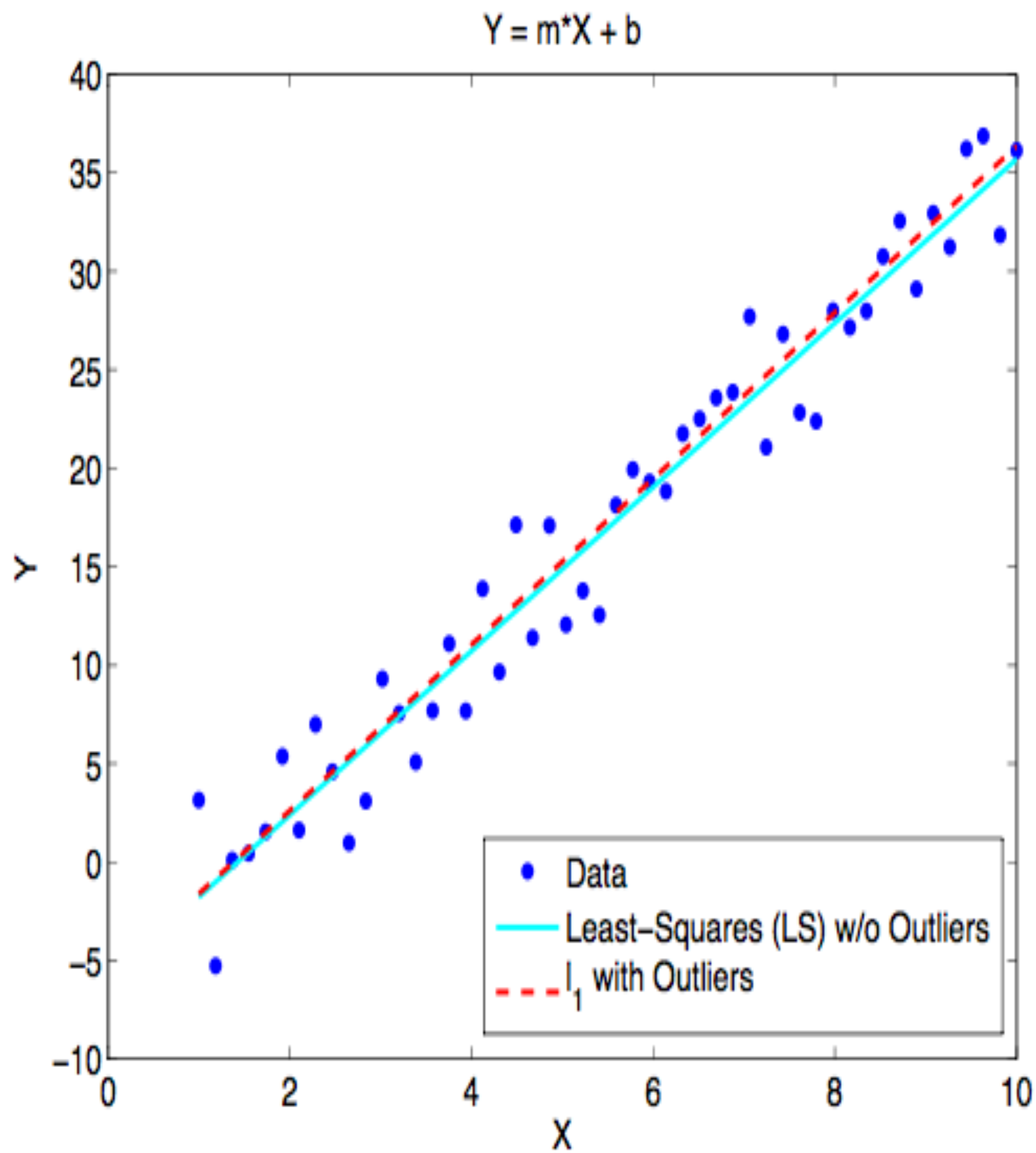
- Include OUTLIERS

- Compare results using different cost functions



Robust Least - Squares

(robust to outliers)

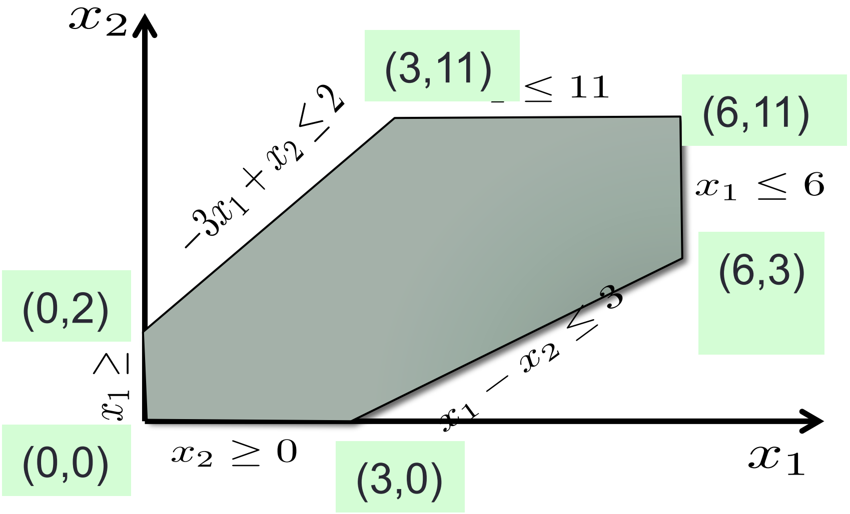


INTERIOR POINT METHODS: BASIC INTRODUCTION

Linear Programming Problem

$$
\begin{array}{rrrcr}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}
$$

Note: Not drawn to scale



Karush-Kuhn-Tucker Conditions

- Very important for many optimization problems.

$$\begin{array}{llcl} \max & \mathbf{c}^\intercal \mathbf{x} & & \\ & A\,\mathbf{x} + \mathbf{x_s} & = & \mathbf{b} \\ & \mathbf{x}, \mathbf{x_s} & \geq & \mathbf{0} \end{array}$$

Primal

$$\begin{array}{llcl} \min & \mathbf{b}^\intercal \mathbf{y} & & \\ & A^\intercal\,\mathbf{y} - \mathbf{y_s} & = & \mathbf{c} \\ & \mathbf{y}, \mathbf{y_s} & \geq & \mathbf{0} \end{array}$$

Dual

Necessary and Sufficient Conditions for optimal solution

$$(\mathbf{x}, \mathbf{x_s}, \mathbf{y}, \mathbf{y_s})$$

KKT conditions for Linear Programs

The primal-dual solution $(\mathbf{x}, \mathbf{x_s}, \mathbf{y}, \mathbf{y_s})$ is optimal iff it satisfies the following conditions:

$$\boxed{\begin{aligned} A\,\mathbf{x} + \mathbf{x_s} &= \mathbf{b} \\ \mathbf{x}, \mathbf{x_s} &\geq \mathbf{0} \end{aligned}}$$

(x,x_s) is primal feasible

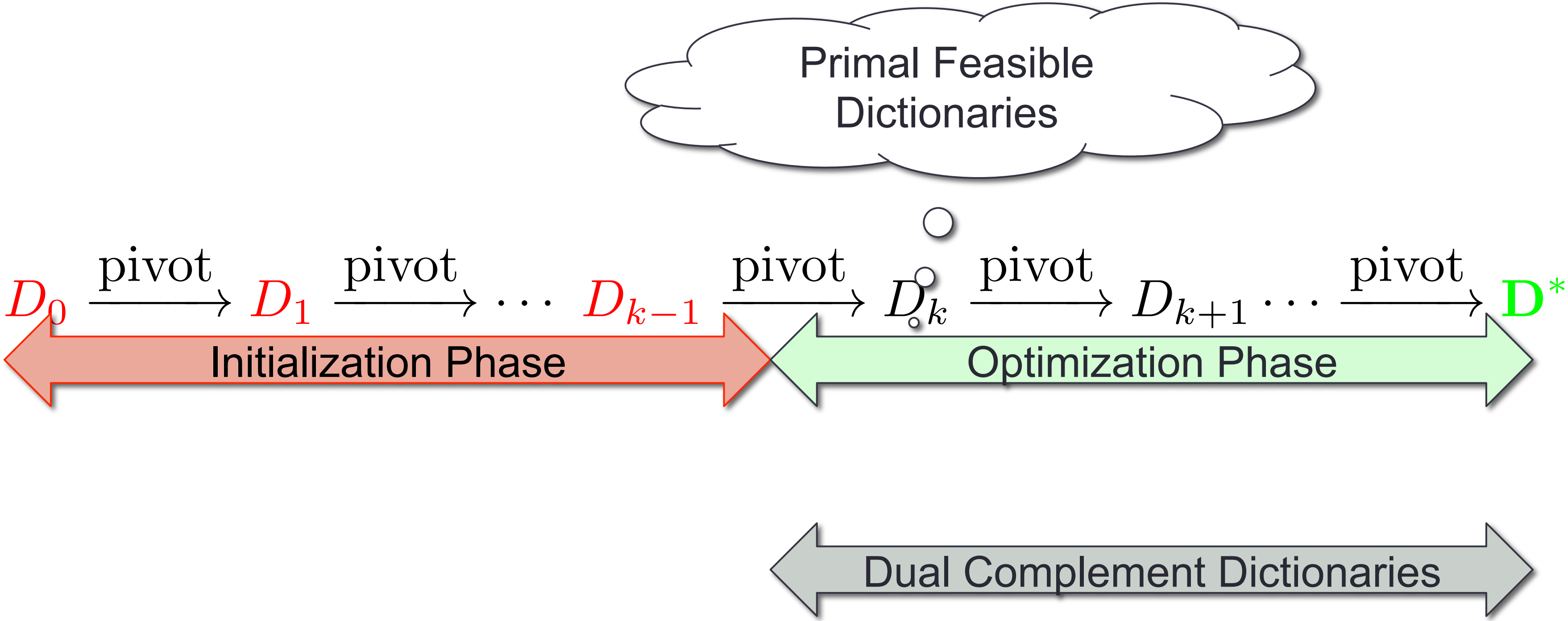
$$\boxed{\begin{aligned} A^\intercal\,\mathbf{y} - \mathbf{y_s} &= \mathbf{c} \\ \mathbf{y}, \mathbf{y_s} &\geq \mathbf{0} \end{aligned}}$$

(y,y_s) is dual feasible

$$\boxed{\begin{aligned} x_j y_{s,j} &= 0 \\ y_j x_{s,j} &= 0 \end{aligned}}$$

Product of complementary pairs is zero.

Simplex Method: Overview



Simplex Method

- Sequence of primal dual solutions (dictionaries)

$$\boxed{(\mathbf{x_0}, \mathbf{y_0}) \ \rightarrow \ (\mathbf{x_1}, \mathbf{y_1}) \rightarrow \ \cdots \ \rightarrow \ (\mathbf{x}^*, \mathbf{y}^*)}$$

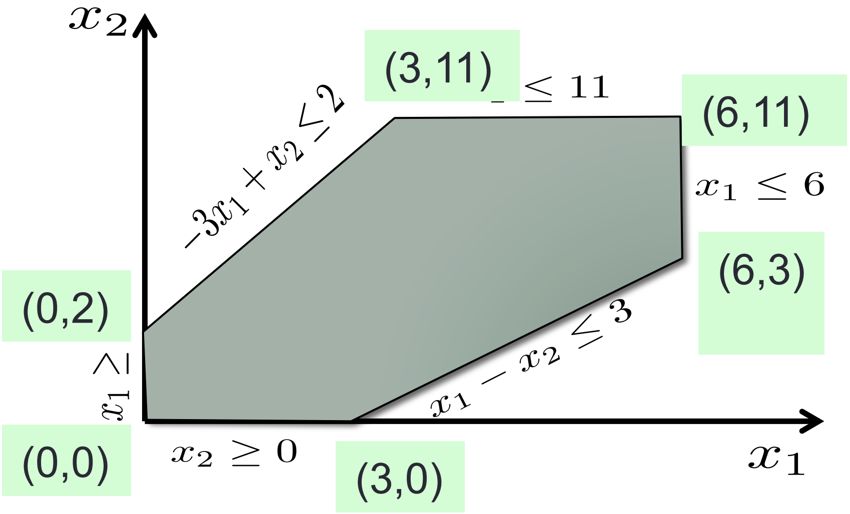
- Maintain Primal Feasibility.
- Maintain Complementarity Conditions.
- Solutions are vertices of primal/dual feasible regions.
- Dual Feasibility achieved only at the very end.

Interior Point Methods

A class of methods.

- Central Path methods
- Affine Scaling Method
- Active Set
- …

Note: Not drawn to scale



Converges to solution vs. Find the precise answer.

Interior Point Methods

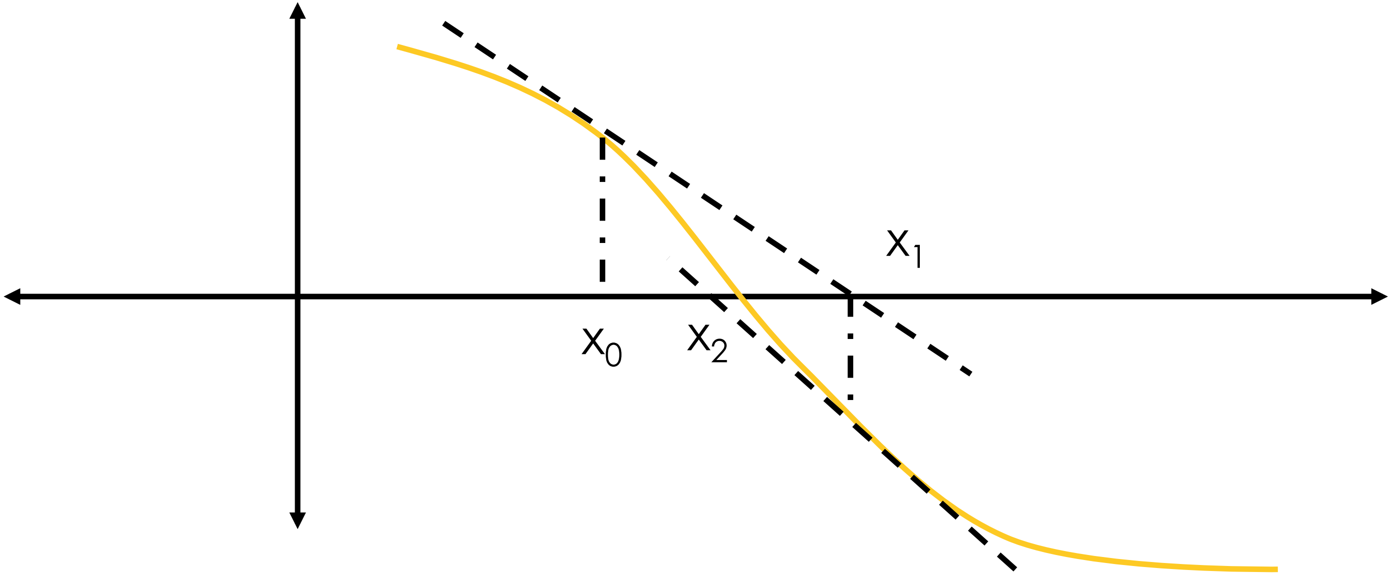
- We will consider a simple central path method.
- Our presentation sequence:
  - Newton’s Method for Solving Equations.
  - Relaxed (mu) complementarity conditions.
  - Central Path.
  - Computing the Newton Step.
  - Adjusting Step Size.
  - Some experiments.

NEWTON'S METHOD

Basic Goal

Solve the (system) of equations: $F(\mathbf{x}) = 0$

$F(\mathbf{x})$ assumed continuous and differentiable (smooth)



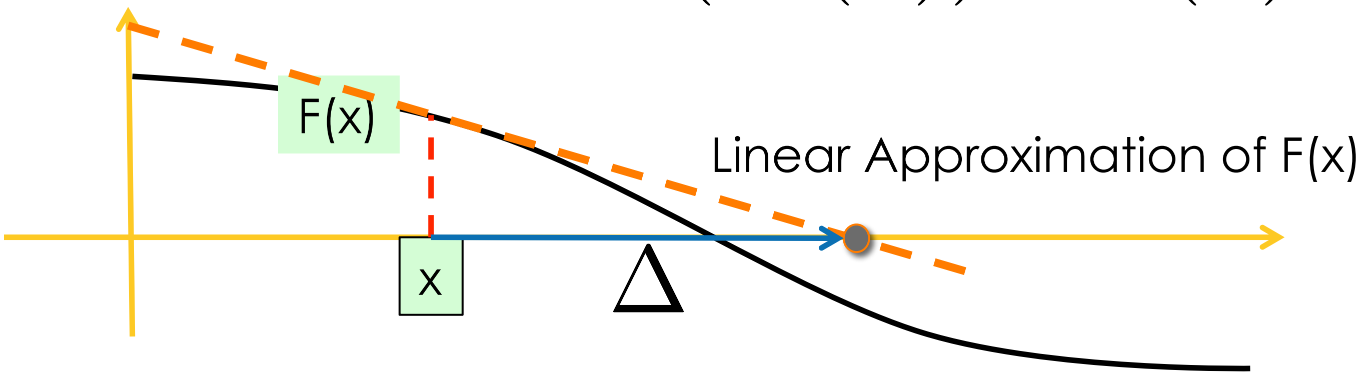
Newton Step

Currently, $F(\mathbf{x}) \neq 0$

Linear Approximation: $F(\mathbf{x}+\Delta) \simeq \; F(\mathbf{x}) + F'(\mathbf{x})\Delta$

$$F(\mathbf{x}) + F'(\mathbf{x})\Delta = \mathbf{0}$$

$$\Delta = -(F'(\mathbf{x}))^{-1}F(\mathbf{x})$$



Newton Step (n-dimensions)

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

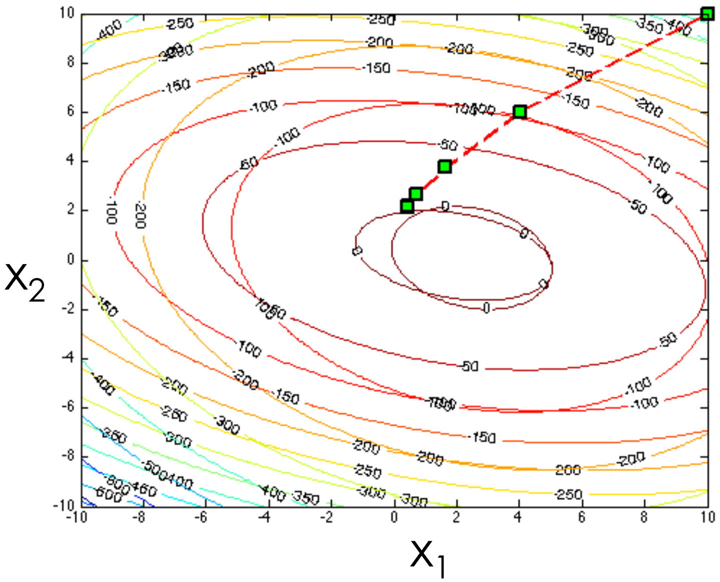
$$F(x) = \begin{bmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{bmatrix}$$

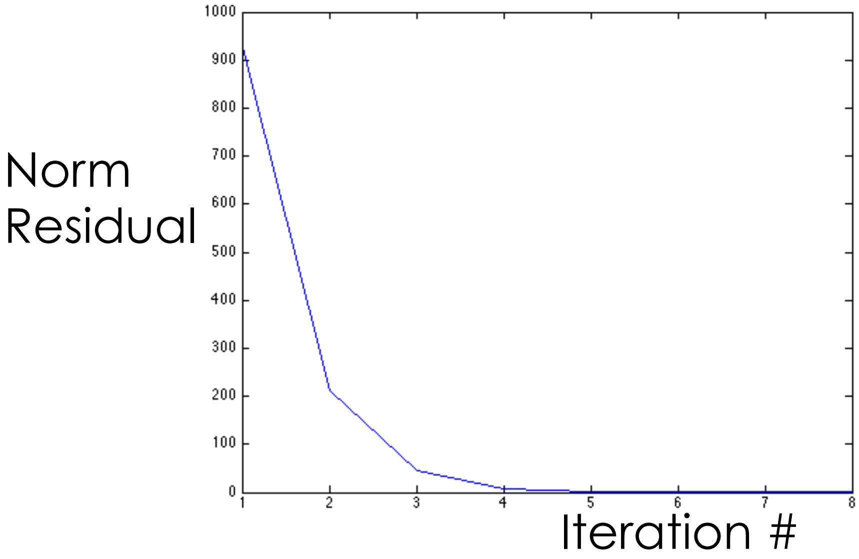
$$F'(\mathbf{x}) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$$

$$\Delta = -(F'(\mathbf{x}))^{-1} F(\mathbf{x})$$

Newton's method example

$$f(x_1, x_2) = \begin{pmatrix} -x_1^2 - 3x_2^2 - x_1 x_2 + 3x_2 + 4x_1 + 5 \\ -2x_1^2 - 3x_2^2 - x_1 x_2 + 10x_1 + 3x_2 \end{pmatrix} = 0$$





Newton Method Convergence

- If method converges, it does so to a root of F(x).

- Convergence is not guaranteed.
  - Only if starting point in the "basin of attraction" of the root.
  - F'(x) should not vanish at the root ("simple" root).

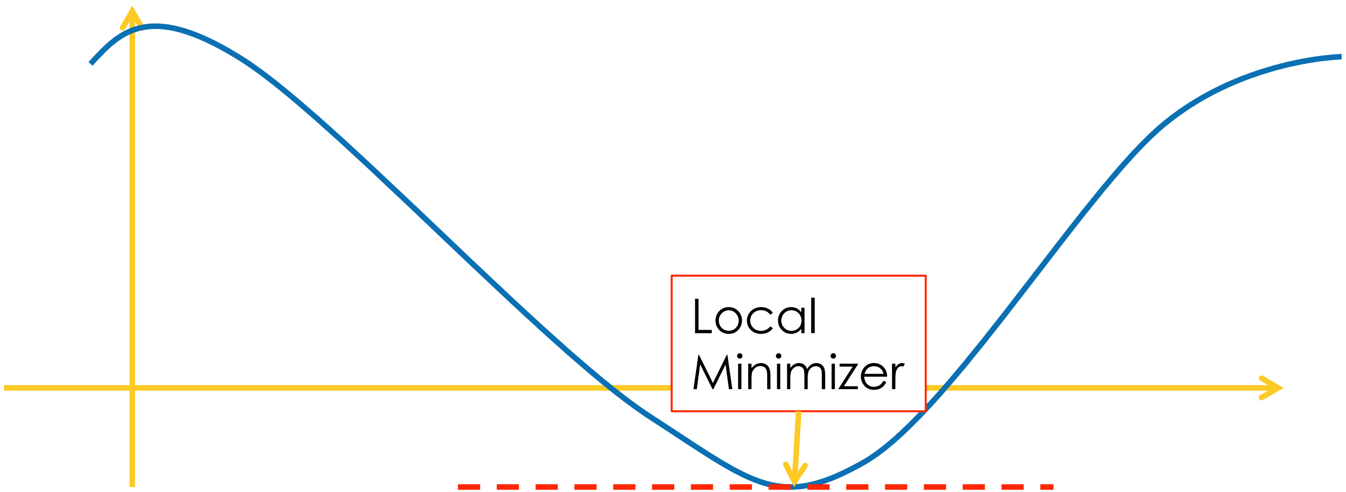
- Convergence is quadratic.
  - Often very fast when it does converge.

Complexity of each newton step.

- Goal: Compute root of F(x) using Newton’s method.
- Compute the Jacobian F’(x)
- Newton Step:
  - Invert the Jacobian and multiply with value of function.

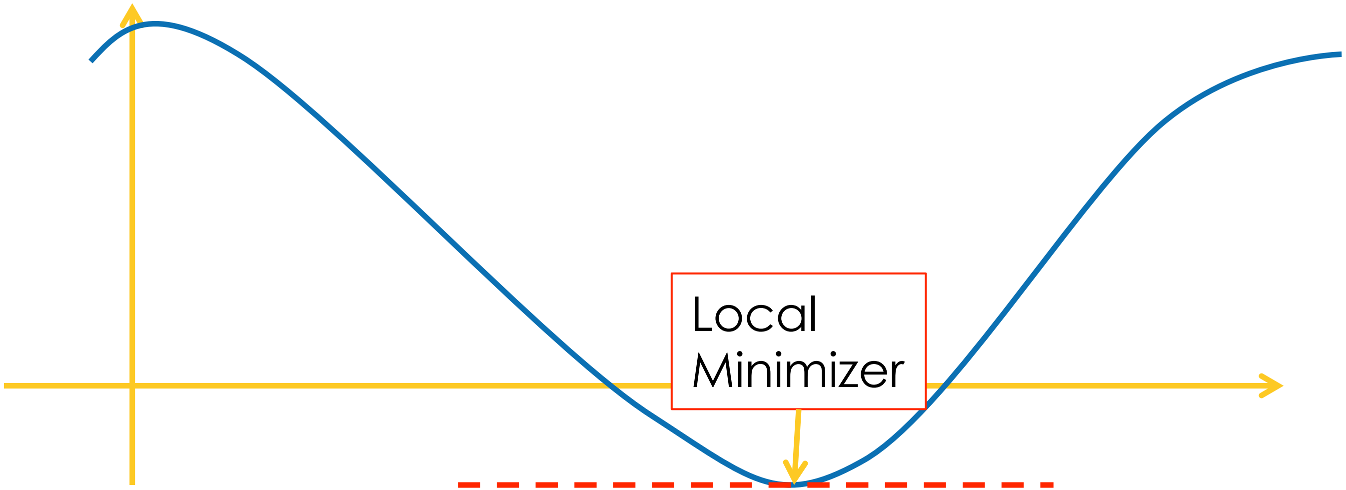
Newton Method for Optimization

- Goal: minimize function F(x) for all x.
- Unconstrained minimization problem.



Newton Method for Optimization

- Goal: minimize function F(x) for all x.
- Unconstrained minimization problem.



Minimization of smooth function.

$$F : \mathbb{R}^n \to \mathbb{R}$$

- F is a C^2 function.
- Continuous, first and second derivatives.

$$\text{If } \mathbf{x} \in \mathbb{R}^n \text{ is a local minimizer of } F \text{ then } \nabla F = 0$$

First-Order Necessary Conditions

$$\text{If } \nabla F(\mathbf{x}) = 0 \text{ and } \nabla^2 F \text{ is positive definite at } \mathbf{x},$$
$$\text{then } \mathbf{x} \text{ is an isolated local minimum of } F.$$

Second-Order Sufficient Condition

Newton method for finding minima

$$F : \mathbb{R}^n \to \mathbb{R}$$

$$\boxed{\nabla F(\mathbf{x}) = \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{pmatrix} = 0}$$

Solve

$$\text{Newton Step: } \Delta = \boxed{-(\nabla^2 F)^{-1}}(\nabla F)$$

Hessian Matrix

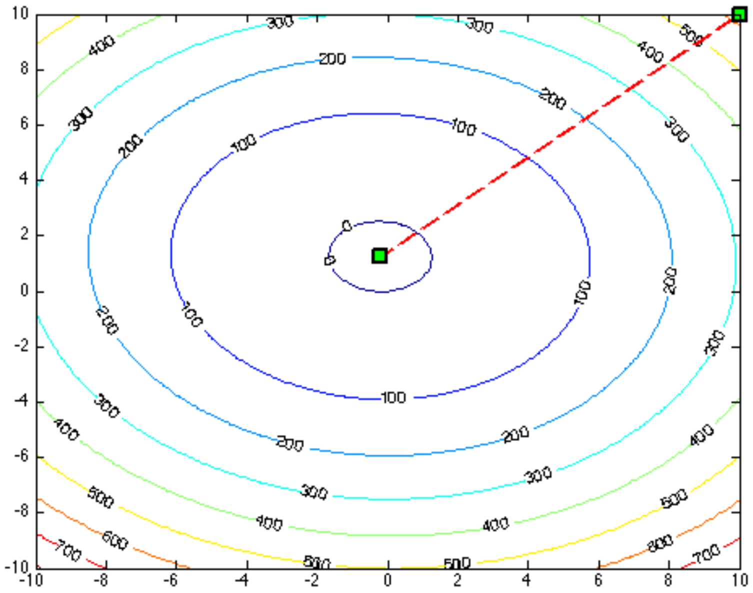
Hessian Matrix

$$\nabla^2 F = \begin{bmatrix} \frac{\partial^2 F}{\partial \ x_1^2} & \frac{\partial^2 F}{\partial \ x_1 x_2} & \cdots & \frac{\partial^2 F}{\partial \ x_1 x_n} \\ \frac{\partial^2 F}{\partial \ x_2 x_1} & \frac{\partial^2 F}{\partial \ x_2^2} & \cdots & \frac{\partial^2 F}{\partial \ x_2 x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 F}{\partial \ x_n x_1} & \frac{\partial^2 F}{\partial \ x_n x_2} & \cdots & \frac{\partial^2 F}{\partial \ x_n^2} \end{bmatrix}$$

Does inverse always exist?

$$\text{Newton Step: } \Delta = -(\nabla^2 F)^{-1}(\nabla F)$$

Newton's Method Example



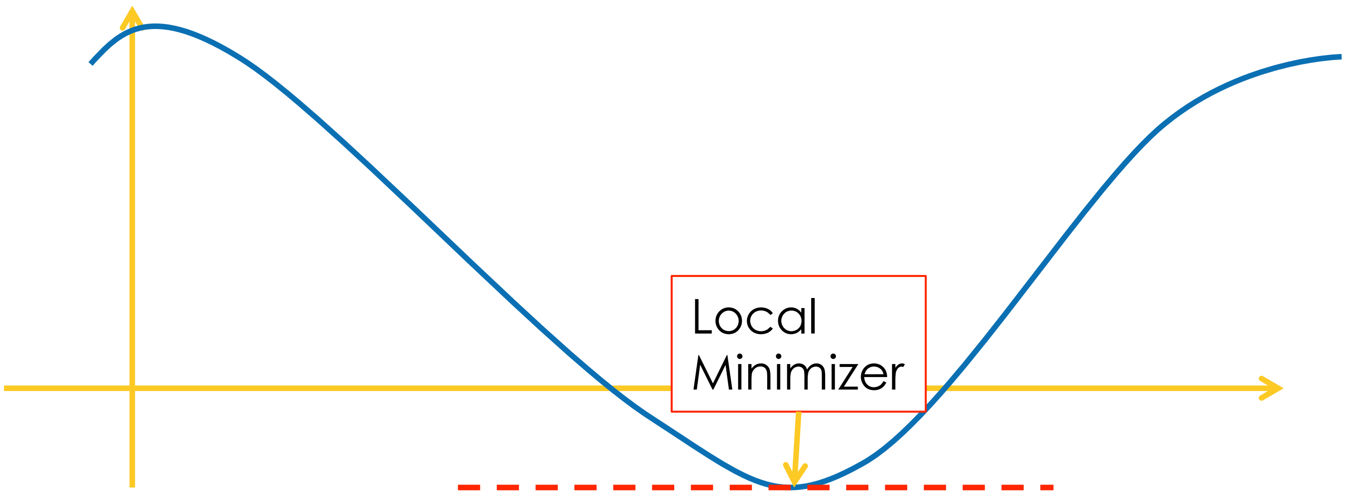
$$\min_{(x,y)}(3x^2 + 4y^2 + 0.2xy + x - 10y)$$

EQUALITY CONSTRAINED OPTIMIZATION

Lagrange Multiplier Method

Unconstrained Optimization

- Goal: minimize function F(x) for all x.
- Unconstrained minimization problem.



Equality Constrained Optimization

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & g_1(\mathbf{x}) = 0 \\ & g_2(\mathbf{x}) = 0 \\ & \vdots \\ & g_m(\mathbf{x}) = 0 \end{array}$$

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \sum_{i=1}^{m} y_i g_i(\mathbf{x})$$

$$\boxed{\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}) &= 0 \\ \nabla_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) &= 0 \end{aligned}}$$

First order Necessary Conditions

Lagrange Multiplier Method

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_1(\mathbf{x}) = 0 \\ & g_2(\mathbf{x}) = 0 \\ & \vdots \\ & g_m(\mathbf{x}) = 0 \end{aligned}$$

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \sum_{i=1}^{m} y_i g_i(\mathbf{x})$$

$$\boxed{\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}) &= 0 \\ \nabla_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) &= 0 \end{aligned}}$$

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j} = 0$$

$$g_i(\mathbf{x}) = 0$$

Solve using Newton's method

Example

$$\min \; \sin(x) + \cos(y) + z^2$$
$$\text{s.t. } x^2 + y^2 + z^2 = 1$$

$$L(x, y, z, \lambda) = \sin(x) + \cos(y) + z^2 + \lambda(x^2 + y^2 + z^2 - 1)$$

$$\begin{aligned} \cos(x) + 2\lambda x &= 0 \\ -\sin(y) + 2\lambda y &= 0 \\ 2z + 2\lambda z &= 0 \\ x^2 + y^2 + z^2 &= 1 \end{aligned}$$

Solve using Newton's method

LOG BARRIER METHOD.

Linear Programming Formulation

$$\begin{aligned} \max \quad & \mathbf{c}^\intercal \mathbf{x} \\ A\mathbf{x} & \leq \mathbf{b} \\ \mathbf{x} & \geq 0 \end{aligned}$$

$$\begin{aligned} \max \quad & \mathbf{c}^\intercal \mathbf{x} \\ A\mathbf{x} + \mathbf{x}_s & = \mathbf{b} \\ \mathbf{x}, \mathbf{x_s} & \geq 0 \end{aligned}$$

Primal Standard form with
Slack Variables

Log Barrier Trick

Inequality constrained optimization:

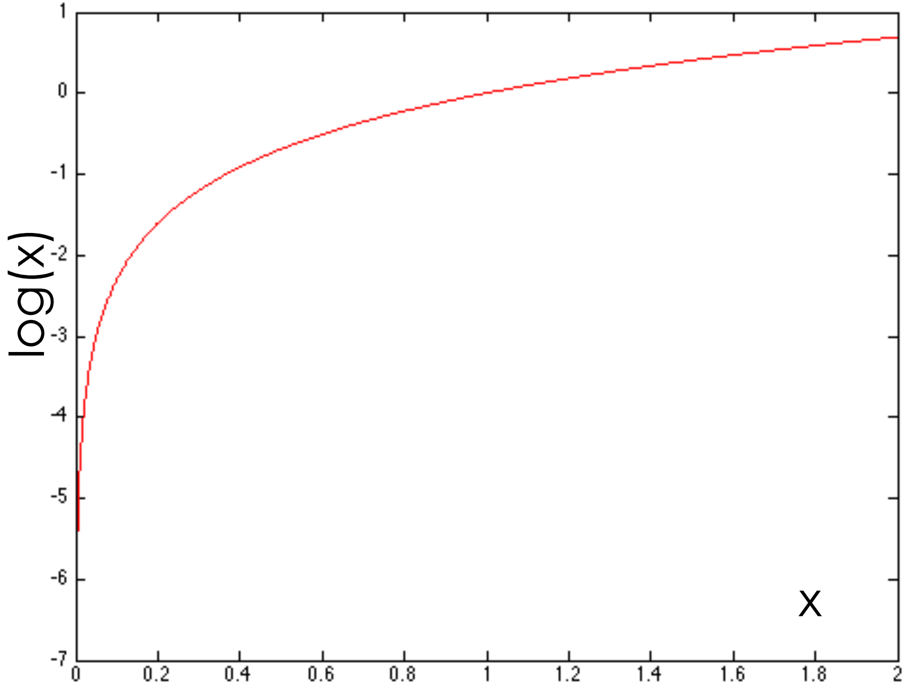
$$\max \; f(x) \text{ s.t. } g(x) \geq 0$$

Log Barrier Transformation of Inequality:

$$\max f(x) + \mu(\log(g(x)))$$

Log Barrier Trick (Log Function)

- log(x) is –Inf if x <= 0
- Adding log(x) to objective forbids x <= 0



Log Barrier Trick

$$\max f(x) + \mu(\log(g(x)))$$

$$\boxed{\text{A } \mu \to 0, \text{ we converge to solution of original problem.}}$$

- Solve log-barrier problem for initial mu (start with g(x) > 0)
- Gradually decrease mu ($\mu \to 0$).
- Stopping criterion: Change in x is below tolerance.

Linear Programming Formulation

$$\begin{aligned} \max \quad & \mathbf{c}^\intercal \mathbf{x} \\ A\mathbf{x} & \leq \mathbf{b} \\ \mathbf{x} & \geq 0 \end{aligned}$$

$$\begin{aligned} \max \quad & \mathbf{c}^\intercal \mathbf{x} \\ A\mathbf{x} + \mathbf{x}_s & = \mathbf{b} \\ \mathbf{x}, \mathbf{x_s} & \geq 0 \end{aligned}$$

Primal Standard form with Slack Variables

Log Barrier Formulation

$$\begin{array}{lrcl} \max & \mathbf{c}^\intercal\ \mathbf{x} & & \\ & A\mathbf{x} + \mathbf{x}_s & = & \mathbf{b} \\ & \mathbf{x}, \mathbf{x_s} & \geq & 0 \end{array}$$

$$\max\ \mathbf{c}^\intercal\ \mathbf{x} + \mu \sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i})$$

$$\text{s.t. } A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$

Equality Constrained Optimization

$$\max \ \mathbf{c}^\intercal \ \mathbf{x} + \mu \textstyle\sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i})$$
$$\text{s.t.} \ A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$

$$\boxed{L(\mathbf{x}, \mathbf{y}) = \left(\begin{array}{c} \mathbf{c}^\intercal \mathbf{x} + \mu \sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i}) \\ + \mathbf{y}^\intercal \ (A\mathbf{x} + \mathbf{x_s} - \mathbf{b}) \end{array} \right)}$$

$$\frac{\partial L}{\partial x_j} = c_j + \frac{\mu}{x_j} + \mathbf{y}^\intercal A_{:,j}$$

MU-COMPLEMENTARITY CONDITIONS

Overview

$$\begin{array}{llcl} \max & \mathbf{c}^\intercal \ \mathbf{x} & & \\ & A\mathbf{x} + \mathbf{x}_s & = & \mathbf{b} \\ & \mathbf{x}, \mathbf{x_s} & \geq & 0 \end{array}$$

Primal Problem

Log Barrier Trick

$$\max \ \mathbf{c}^\intercal \ \mathbf{x} + \mu \textstyle\sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i})$$
$$\text{s.t. } A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$

A $\mu \rightarrow 0$, we converge to solution of original problem.

Lagrange Multiplier Method

Lagrange Multiplier Method

$$L(\mathbf{x},\mathbf{y}) = \left(\begin{array}{c} \mathbf{c}^\intercal \mathbf{x} + \mu \sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i}) \\ +\mathbf{y}^\intercal \left(A\mathbf{x} + \mathbf{x_s} - \mathbf{b} \right) \end{array} \right)$$

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x},\mathbf{y}) &= 0 \\ \nabla_{\mathbf{y}} L(\mathbf{x},\mathbf{y}) &= 0 \end{aligned}$$

First Order Necessary Conditions.

Mu KKT conditions

$$\boxed{\begin{aligned} A\mathbf{x} + \mathbf{x_s} &= \mathbf{b} && \text{Primal} \\ A^\intercal \mathbf{y} - \mathbf{y_s} &= \mathbf{c} && \text{Dual} \\ XY_s\mathbf{e} &= \mu\mathbf{e} && \text{Mu-Complementarity} \\ X_sY\mathbf{e} &= \mu\mathbf{e} && \end{aligned}}$$

$$\begin{aligned} X &= \text{diag}(\mathbf{x}) \\ X_s &= \text{diag}(\mathbf{x_s}) \\ Y &= \text{diag}(\mathbf{y}) \\ Y_s &= \text{diag}(\mathbf{y_s}) \end{aligned}$$

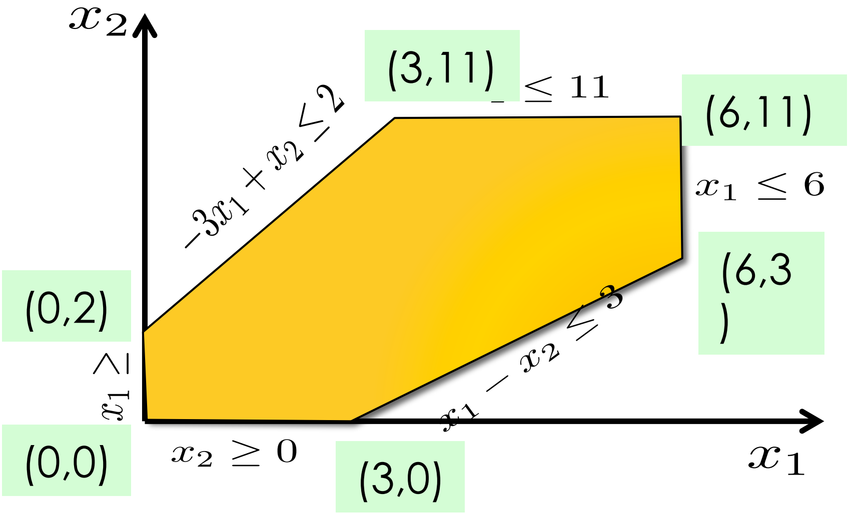
As mu approaches 0,
we obtain
KKT conditions!!

THE CENTRAL PATH

Linear Programming Problem

$$
\begin{array}{rrrcl}
\text{max.} & x_1 & +2x_2 & & \\
\text{s.t.} & -3x_1 & +x_2 & \leq & 2 \\
 & & +x_2 & \leq & 11 \\
 & x_1 & -x_2 & \leq & 3 \\
 & x_1 & & \leq & 6 \\
 & x_1, & x_2 & \geq & 0
\end{array}
$$

Note: Not drawn to scale



Overview

$$\begin{array}{lrcl} \max & \mathbf{c}^\intercal\ \mathbf{x} & & \\ & A\mathbf{x} + \mathbf{x}_s & = & \mathbf{b} \\ & \mathbf{x}, \mathbf{x_s} & \geq & 0 \end{array}$$

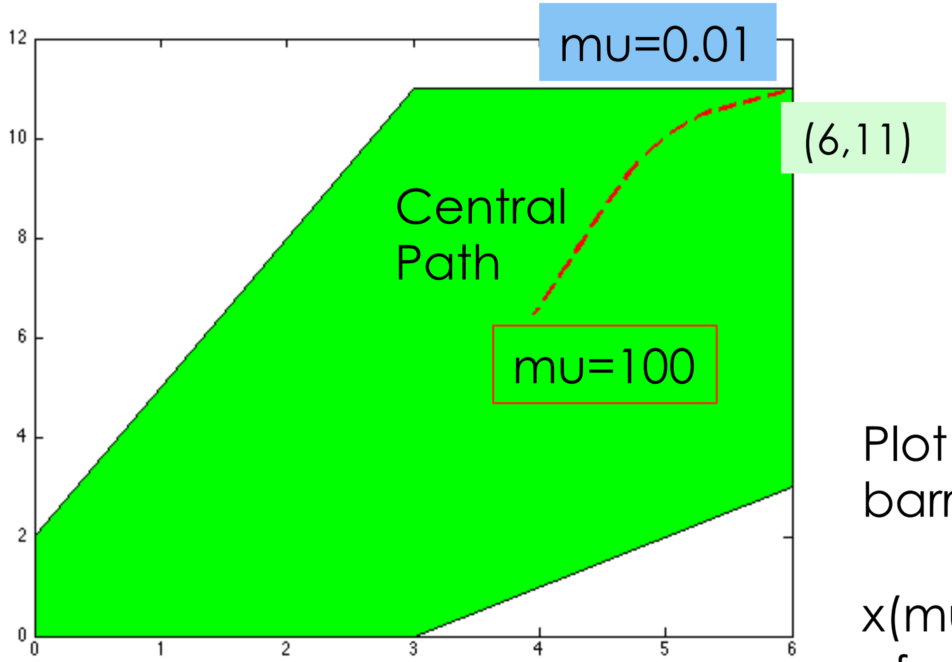
Primal Problem

Log Barrier Trick

$$\max\ \mathbf{c}^\intercal\ \mathbf{x} + \mu \sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(x_{s,i})$$
$$\text{s.t. } A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$

A $\mu \to 0$, we converge to solution of original problem.

Central Path



Plot optimal solutions for barrier problem:

x(mu) as a function of mu.

Solving Linear Programs

- Start with a large value of mu.
  - Use Newton's method to solve for mu-KKT conditions.
- As we iterate, gradually reduce mu.
  - mu' = 0.1 * mu
- Stop when value of primal infeasibility, dual infeasibility, and mu are small enough.

APPLYING NEWTON METHOD TO SOLVING LPS

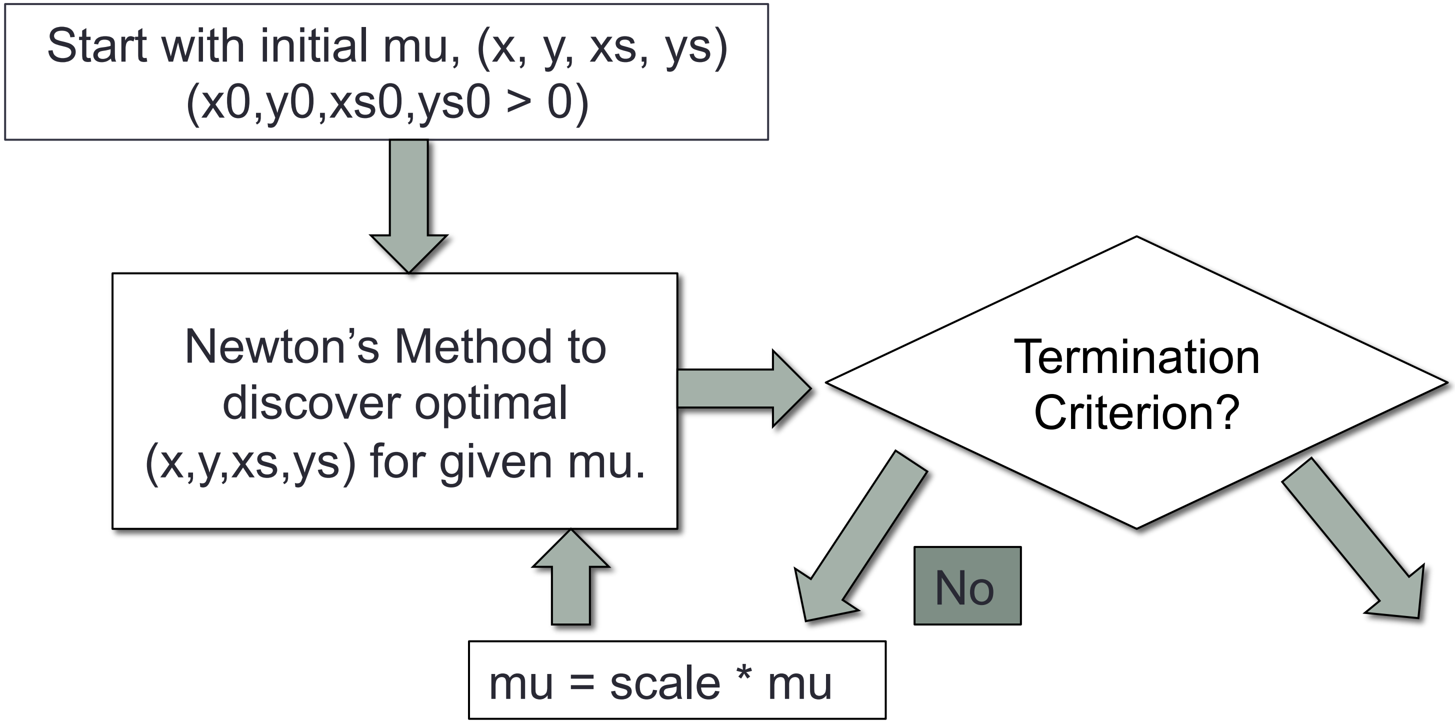
Mu KKT conditions

$$\boxed{\begin{aligned} A\mathbf{x} + \mathbf{x_s} &= \mathbf{b} && \text{Primal} \\ A^\intercal \mathbf{y} - \mathbf{y_s} &= \mathbf{c} && \text{Dual} \\ XY_s\mathbf{e} &= \mu\mathbf{e} && \text{Mu-Complementarity} \\ X_sY\mathbf{e} &= \mu\mathbf{e} && \end{aligned}}$$

$$\begin{aligned} X &= \text{diag}(\mathbf{x}) \\ X_s &= \text{diag}(\mathbf{x_s}) \\ Y &= \text{diag}(\mathbf{y}) \\ Y_s &= \text{diag}(\mathbf{y_s}) \end{aligned}$$

As mu approaches 0,
we obtain
KKT conditions!!

Overall Algorithm



Newton Step

$$A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$ Primal

$$A^\intercal \mathbf{y} - \mathbf{y_s} = \mathbf{c}$$ Dual

$$XY_s\mathbf{e} = \mu\mathbf{e}$$ Mu-Complementarity

$$X_sY\mathbf{e} = \mu\mathbf{e}$$

Solve for $F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s) = \mathbf{0}$

$$\boxed{F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y_s}) = \begin{bmatrix} A\mathbf{x} + \mathbf{x}_s - \mathbf{b} \\ A^\intercal \mathbf{y} = \mathbf{y_s} - \mathbf{c} \\ XY_s\mathbf{e} - \mu\mathbf{e} \\ X_sY\mathbf{e} - \mu\mathbf{e} \end{bmatrix}}$$

Calculating Newton Step -1

$$F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y_s}) = \begin{bmatrix} A\mathbf{x} + \mathbf{x}_s - \mathbf{b} \\ A^{\intercal}\mathbf{y} = \mathbf{y_s} - \mathbf{c} \\ XY_s\mathbf{e} - \mu\mathbf{e} \\ X_sY\mathbf{e} - \mu\mathbf{e} \end{bmatrix}$$

$$\nabla F = \begin{bmatrix} A & I_{m\times m} & 0_{m\times m} & 0_{m\times n} \\ 0_{n\times m} & 0_{n\times m} & A^{\intercal} & -I_{n\times n} \\ Y_s & 0_{n\times m} & 0_{n\times m} & X \\ 0_{m\times n} & Y & X_s & 0_{m\times n} \end{bmatrix}$$

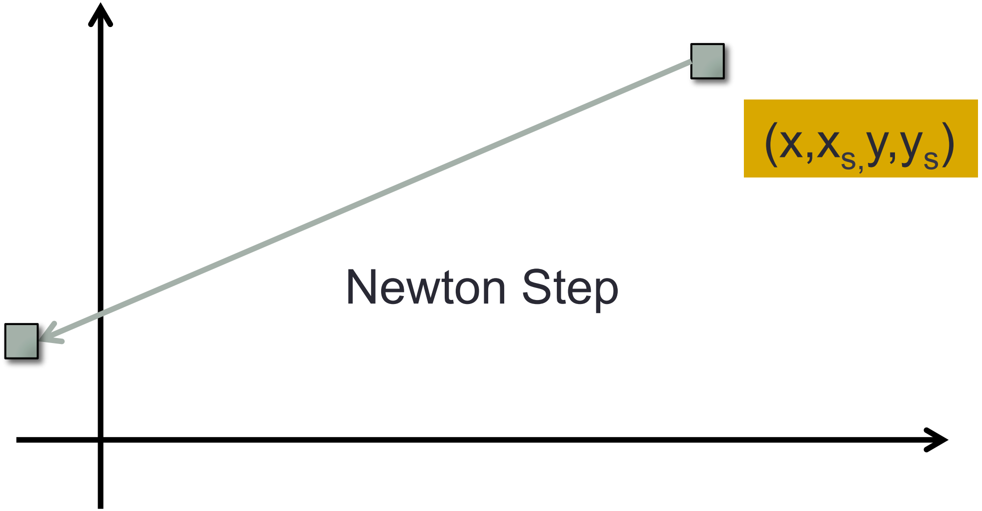
$$\Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s) = -(\nabla F)^{-1} \times F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Calculating Newton Step -2

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Warning: This can violate the non-negativity of (x, x_s, y, y_s)

Sizing the Newton Step



Newton step can make some components of x, x_s, y, y_s negative.

Applying the Newton Step

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

WRONG!!

Use a scale factor $\lambda \in (0,1)$

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \boxed{\lambda} * \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Finding Scale Factor

- Find (*largest)*λsuch that

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \lambda * \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

guarantees that

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} > 0$$

Implementation Detail: We use a smaller value of λ than the largest possible

APPLYING NEWTON METHOD TO SOLVING LPS

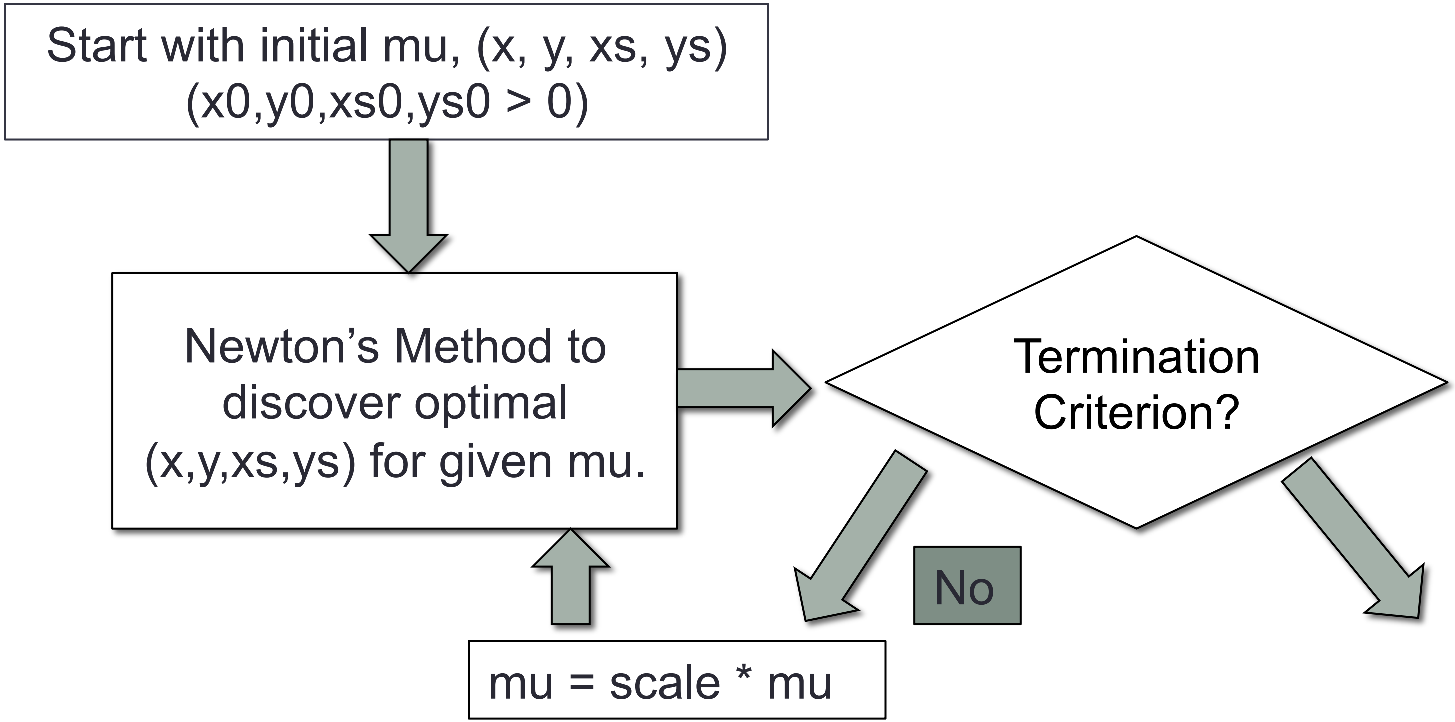
Mu KKT conditions

$$\boxed{\begin{aligned} A\mathbf{x} + \mathbf{x_s} &= \mathbf{b} && \text{Primal} \\ A^\intercal \mathbf{y} - \mathbf{y_s} &= \mathbf{c} && \text{Dual} \\ XY_s\mathbf{e} &= \mu\mathbf{e} && \text{Mu-Complementarity} \\ X_sY\mathbf{e} &= \mu\mathbf{e} && \end{aligned}}$$

$$\begin{aligned} X &= \text{diag}(\mathbf{x}) \\ X_s &= \text{diag}(\mathbf{x_s}) \\ Y &= \text{diag}(\mathbf{y}) \\ Y_s &= \text{diag}(\mathbf{y_s}) \end{aligned}$$

As mu approaches 0, we obtain KKT conditions!!

Overall Algorithm



Newton Step

$$A\mathbf{x} + \mathbf{x_s} = \mathbf{b}$$ Primal

$$A^\intercal \mathbf{y} - \mathbf{y_s} = \mathbf{c}$$ Dual

$$XY_s\mathbf{e} = \mu\mathbf{e}$$ Mu-Complementarity

$$X_sY\mathbf{e} = \mu\mathbf{e}$$

Solve for $F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s) = \mathbf{0}$

$$\boxed{F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y_s}) = \begin{bmatrix} A\mathbf{x} + \mathbf{x}_s - \mathbf{b} \\ A^\intercal \mathbf{y} = \mathbf{y_s} - \mathbf{c} \\ XY_s\mathbf{e} - \mu\mathbf{e} \\ X_sY\mathbf{e} - \mu\mathbf{e} \end{bmatrix}}$$

Calculating Newton Step -1

$$F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y_s}) = \begin{bmatrix} A\mathbf{x} + \mathbf{x}_s - \mathbf{b} \\ A^\intercal \mathbf{y} = \mathbf{y_s} - \mathbf{c} \\ XY_s\mathbf{e} - \mu\mathbf{e} \\ X_sY\mathbf{e} - \mu\mathbf{e} \end{bmatrix}$$

$$\nabla F = \begin{bmatrix} A & I_{m\times m} & 0_{m\times m} & 0_{m\times n} \\ 0_{n\times m} & 0_{n\times m} & A^\intercal & -I_{n\times n} \\ Y_s & 0_{n\times m} & 0_{n\times m} & X \\ 0_{m\times n} & Y & X_s & 0_{m\times n} \end{bmatrix}$$

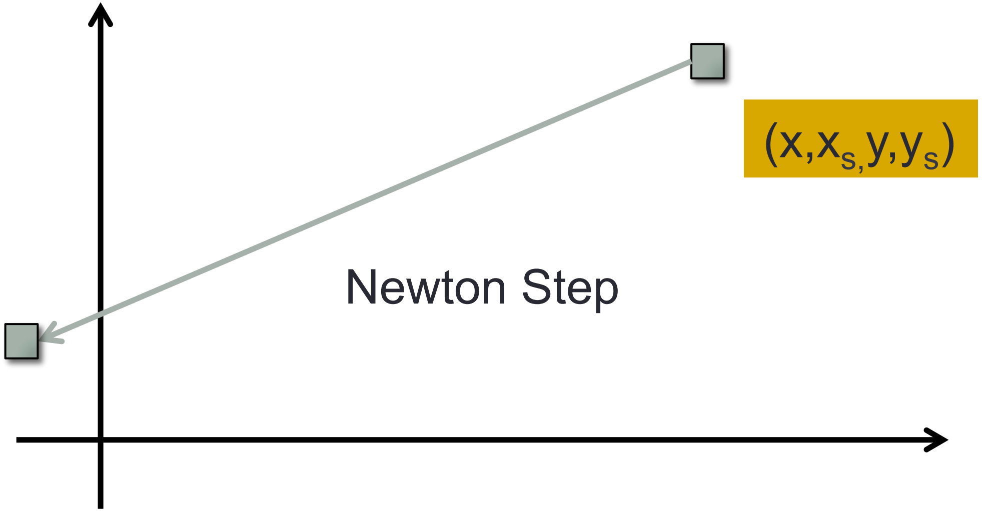
$$\Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s) = -(\nabla F)^{-1} \times F(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Calculating Newton Step -2

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Warning: This can violate the non-negativity of (x, x_s, y, y_s)

Sizing the Newton Step



Newton step can make some components of x,x_s,y,y_s negative.

Applying the Newton Step

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

WRONG!!

Use a scale factor λ ∈ (0,1)

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \lambda * \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

Finding Scale Factor

- Find (*<u>largest)</u>*λsuch that

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} := \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} + \lambda * \Delta(\mathbf{x}, \mathbf{x}_s, \mathbf{y}, \mathbf{y}_s)$$

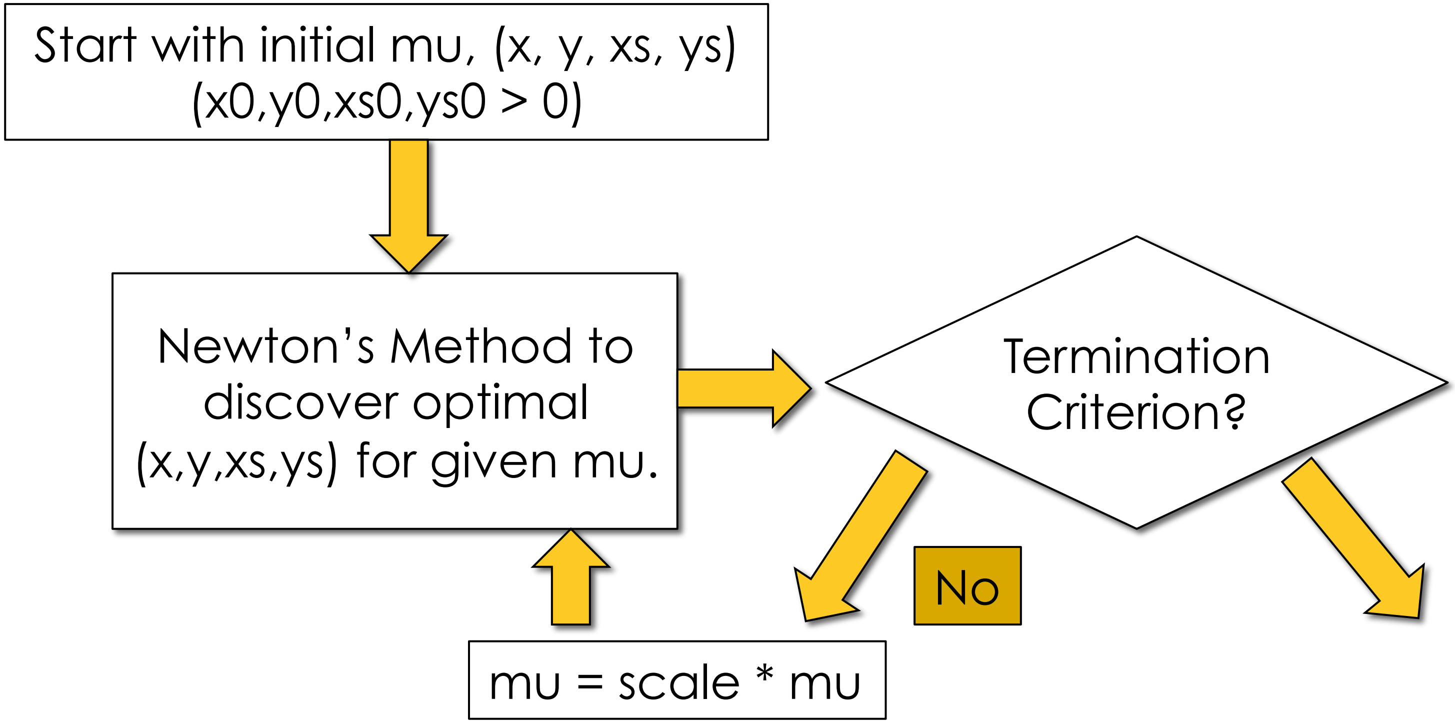
guarantees that

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_s \\ \mathbf{y} \\ \mathbf{y}_s \end{pmatrix} > 0$$

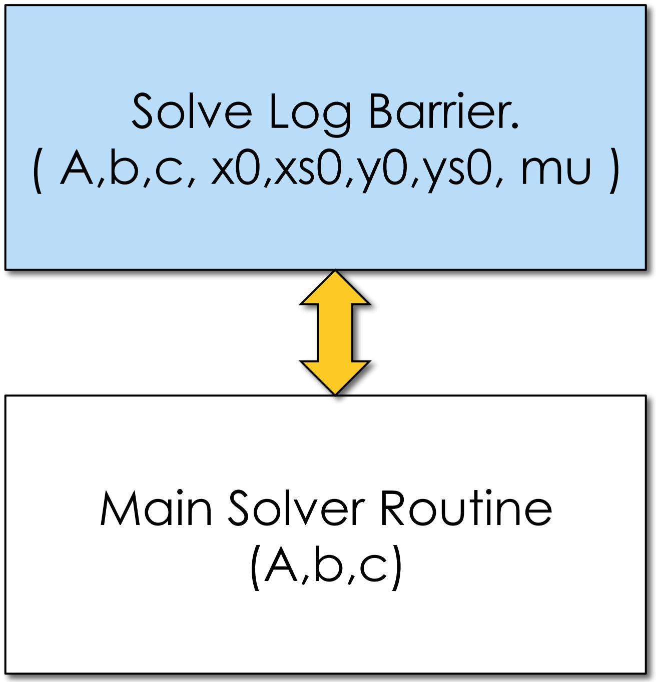
Implementation Detail: We use a smaller value of λ than the largest possible

IMPLEMENTING A SIMPLE INTERIOR POINT SOLVER

Overall Algorithm



Implementation Details



Implement Newton's Method.

Solve to a tolerance or fixed number of iterations.

Implement Termination Criterion.

Analyze the solutions to determine final result.

Termination Criterion

- Primal-Dual Gap: $\mathbf{c}^\intercal\mathbf{x} - \mathbf{b}^\intercal\mathbf{y}$
- Primal Infeasibility Gap: $||A\mathbf{x} + \mathbf{x}_s - \mathbf{b}||$
- Dual Infeasibility Gap: $||A^\intercal\mathbf{y} - \mathbf{y}_s - \mathbf{c}||$
- Value of mu
- Change in solution across iterations.
- Iteration Limit.

Termination with Optimal Value

- The primal-dual gap, primal/dual infeasibilities converge to values less than a tolerance.

```
Optimal solution found with objective value: 1406.699737
Number of iterations to converge: 30
Primal Feasibility Gap: 0.000000
Dual Feasibility Gap: 0.000000
Primal-Dual Gap: 0.000000
KKT-residual: 0.000000  (mu = 0.000000)
Number of Iterations: 30 (LIMIT: 30 )
```

Termination with Primal Unbounded

Dual is infeasible, large values of primal may be seen.

- Primal infeasibility gap may converge to 0, but dual does not.
- Often, Hessian faces condition number issues.
  - Inverting Hessian leads to numerical instabilities.
- Primal-Dual gap does not converge to 0.