Quadratic programs
Cone programs

10-725 Optimization

Geoff Gordon
Ryan Tibshirani

Administrivia

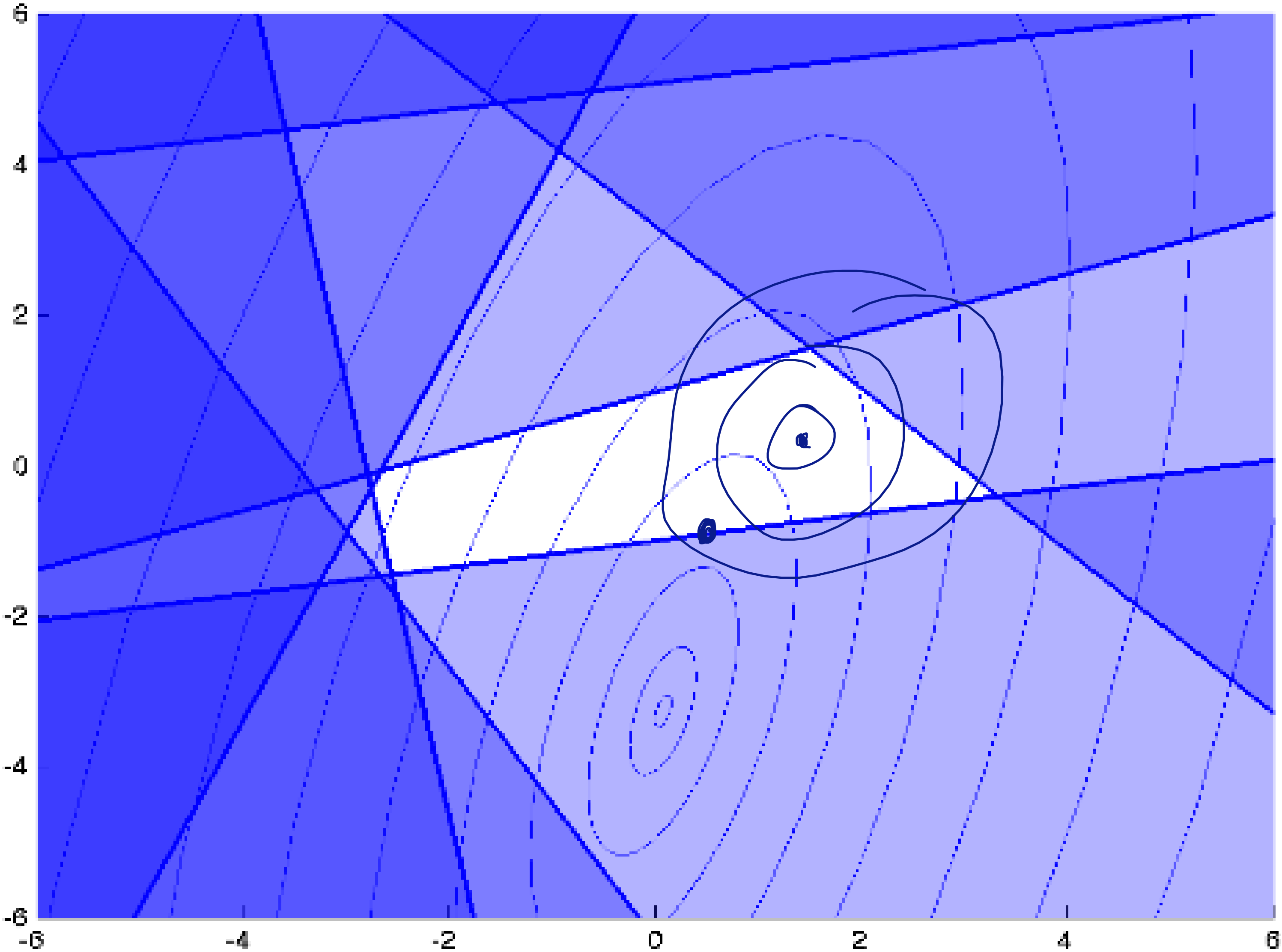
- HW3 back at end of class
- Last day for feedback survey
- All lectures now up on Youtube (and continue to be downloadable from course website)
- Reminder: midterm next Tuesday 11/6!
    - in class, 1 hr 20 min, one sheet (both sides) of notes

Quadratic programs

- m constraints, n vars
  - A: $R^{m \times n}$ b: R^m c: R^n x: R^n H: $R^{n \times n}$
  - [min or max] $x^T H x/2 + c^T x$
  - s.t. $Ax \leq b$ or $Ax = b$ [or some mixture]
  - may have (some elements of) $x \geq 0$
- Convex problem if:
  - min $H \succeq 0$
    max $H \preceq 0$

$$\max\ 2x+x^2+y^2 \text{ s.t.}$$
$$x + y \leq 4$$
$$2x + 5y \leq 12$$
$$x + 2y \leq 5$$
$$x, y \geq 0$$

For example



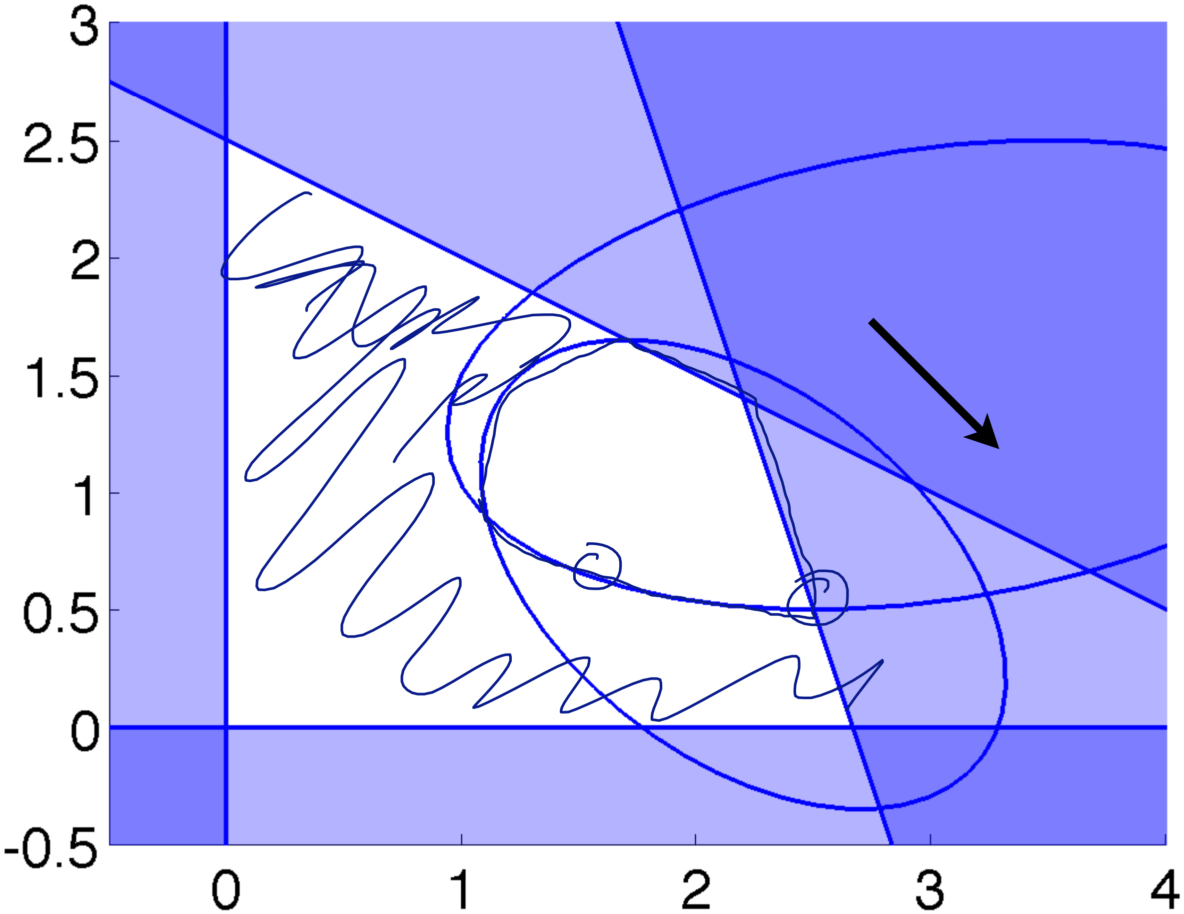
Cone programs

- m constraints, n vars
  - A: $R^{m \times n}$ b: R^m c: R^n x: R^n
  - Cones $K \subseteq R^m$ $L \subseteq R^n$
  - [min or max] $c^T x$ s.t. $\underline{Ax + b \in K}$ $\underline{x \in L}$
  - convex if K, L convex
- E.g., K = $\{0\}^p \times \mathbb{R}_+^q$
- E.g., L = $\{0\}^{p'} \times \mathbb{R}_+^{q'} \times \mathbb{R}^{q''}$

LPs

For example: SOCP

- max
- min $c^T x$ s.t. $A_i x + b_i \in K_i$, $i = 1, 2, \ldots$



$x \in \mathbb{R}^n$

$A_i \in \mathbb{R}^{m_i \times n}$ $b_i \in \mathbb{R}^{m_i}$

K_i $\{0\}$

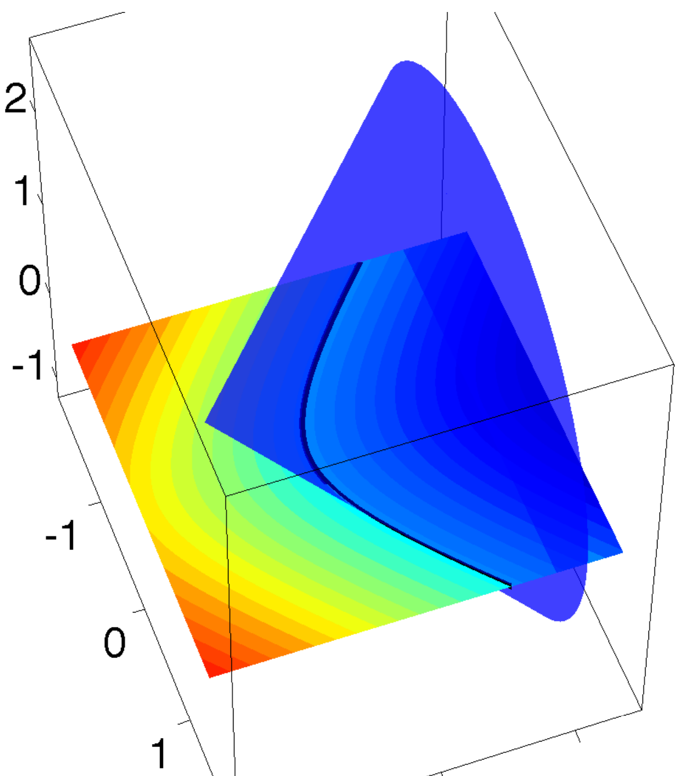
$\mathbb{R}_+$

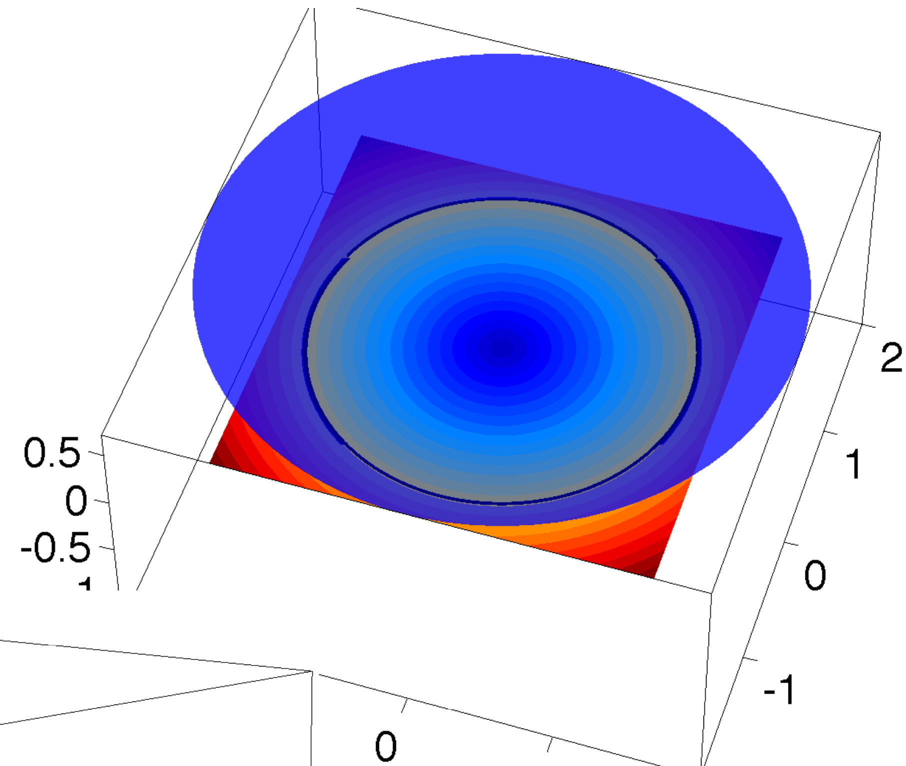
SOC: $\{(y,t) \mid \|y\| \le t\}$

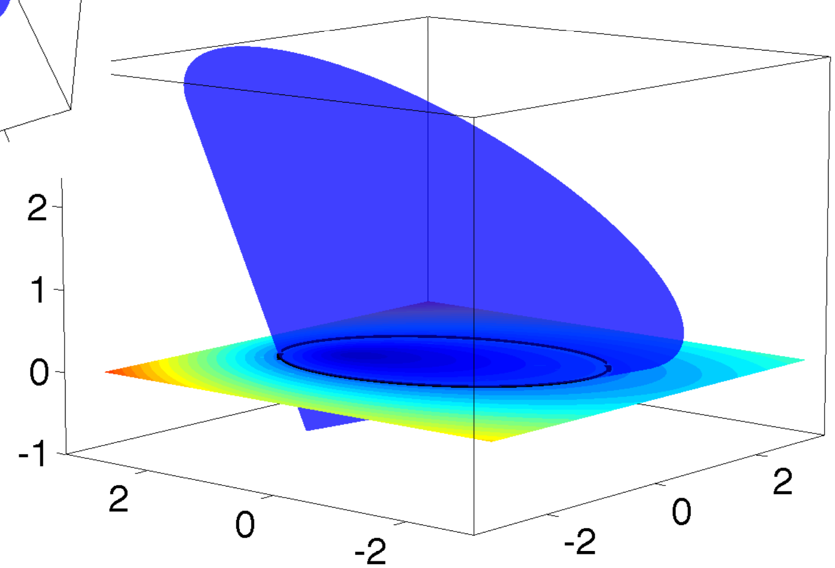
$\mathbb{R}^m$ $\mathbb{R}$

$\mathbb{R}^{m-1}$

Conic sections







QPs are reducible to SOCPs

- min $x^T H x/2 + c^T x$ s.t. …

$H = R^T R$

$\min\ t + c^T x$ s.t. $\underline{t \geq x^T H x/2}$, …

$x^T H x = (x^T R^T)(Rx)$

$= \|Rx\|_2^2$

$\begin{pmatrix} Rx \\ t \\ t+1 \end{pmatrix} \in SOC \qquad t+1 \geq \sqrt{\|Rx\|_2^2 + t^2}$

$\cancel{t^2} + 2t + 1 \geq \|Rx\|_2^2 + \cancel{t^2}$

$t \geq \|Rx\|_2^2/2 - 1/2$

∃ *SOCPs that aren't QPs?*

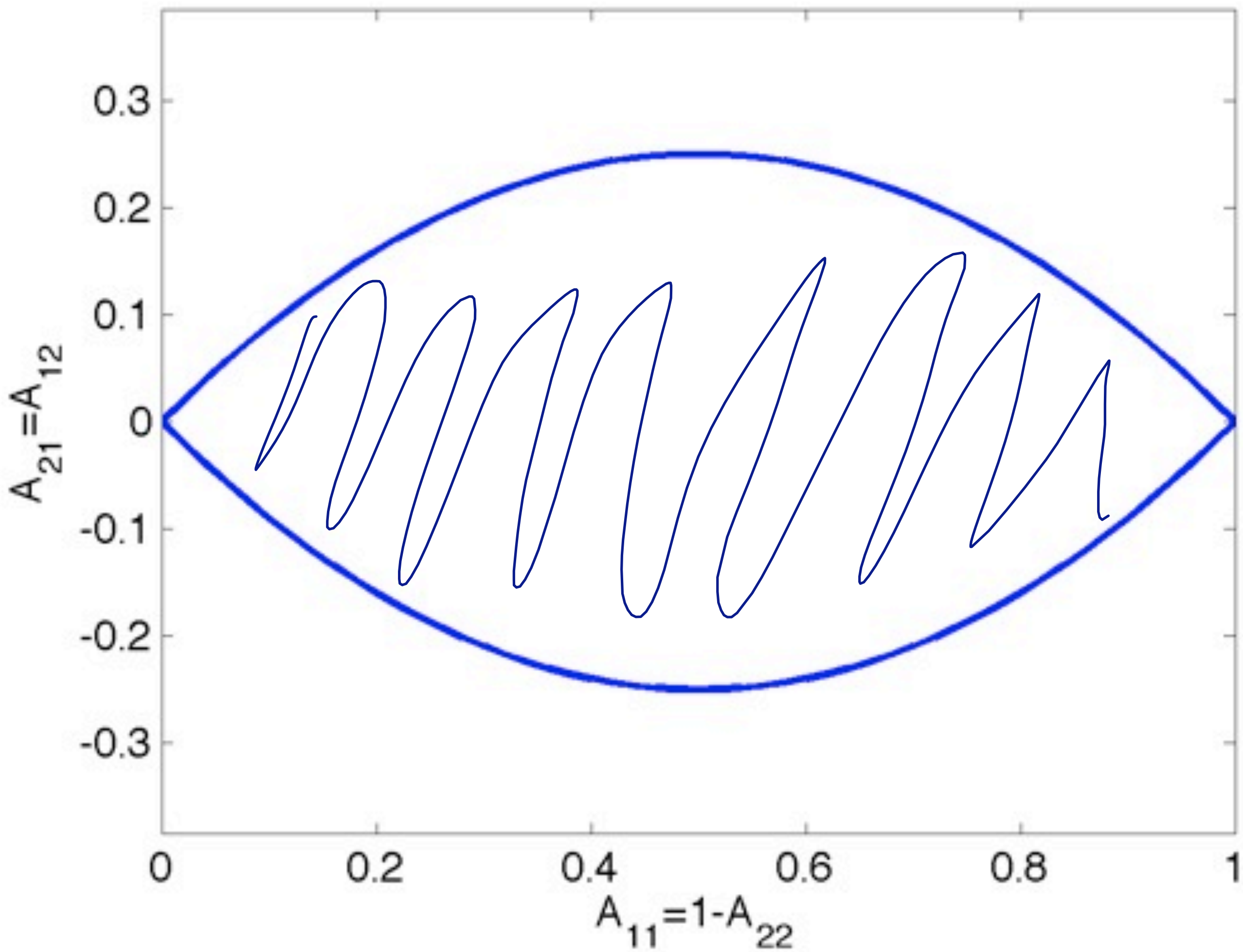
- QCQP: convex quadratic objective & constraints
- minimize $a^2 + b^2$ s.t. min $x^4 + y^4$ s.t. $2x + y = 4$
  - $a \geq x^2$, $b \geq y^2$
  - $2x + y = 4$
- Not a QP (nonlinear constraints)
  - but, can rewrite as SOCP

More cone programs: SDP

- Semidefinite constraint:
  - variable $x \in R^n$
  - constant matrices A_0 $A_1, A_2, \ldots \in R^{m \times m}$
  - constrain $A_0 + \sum_i x_i A_i \succeq 0$ and symmetric $\in S_+^m$
- Semidefinite program: min $c^T x$ s.t.
  - semidefinite constraints
  - linear equalities
  - linear inequalities

Visualizing S_+

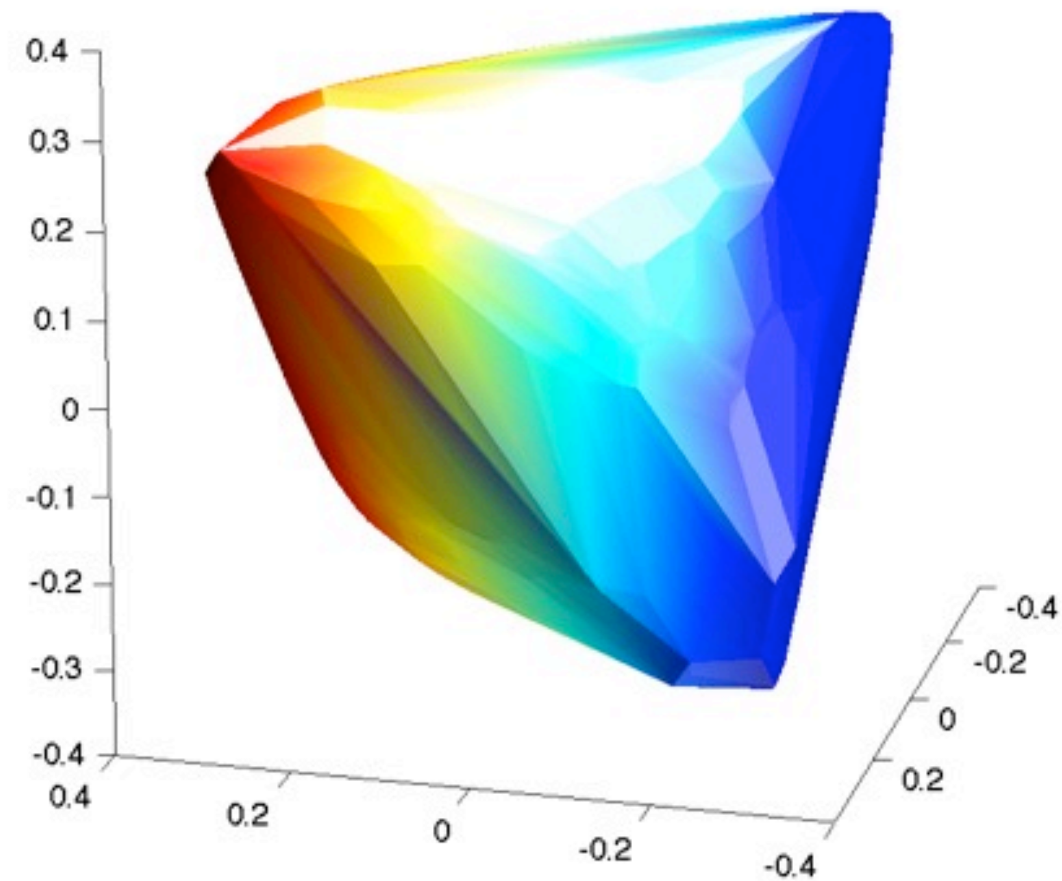
- 2 x 2 symmetric matrices w/ tr(A) = 1



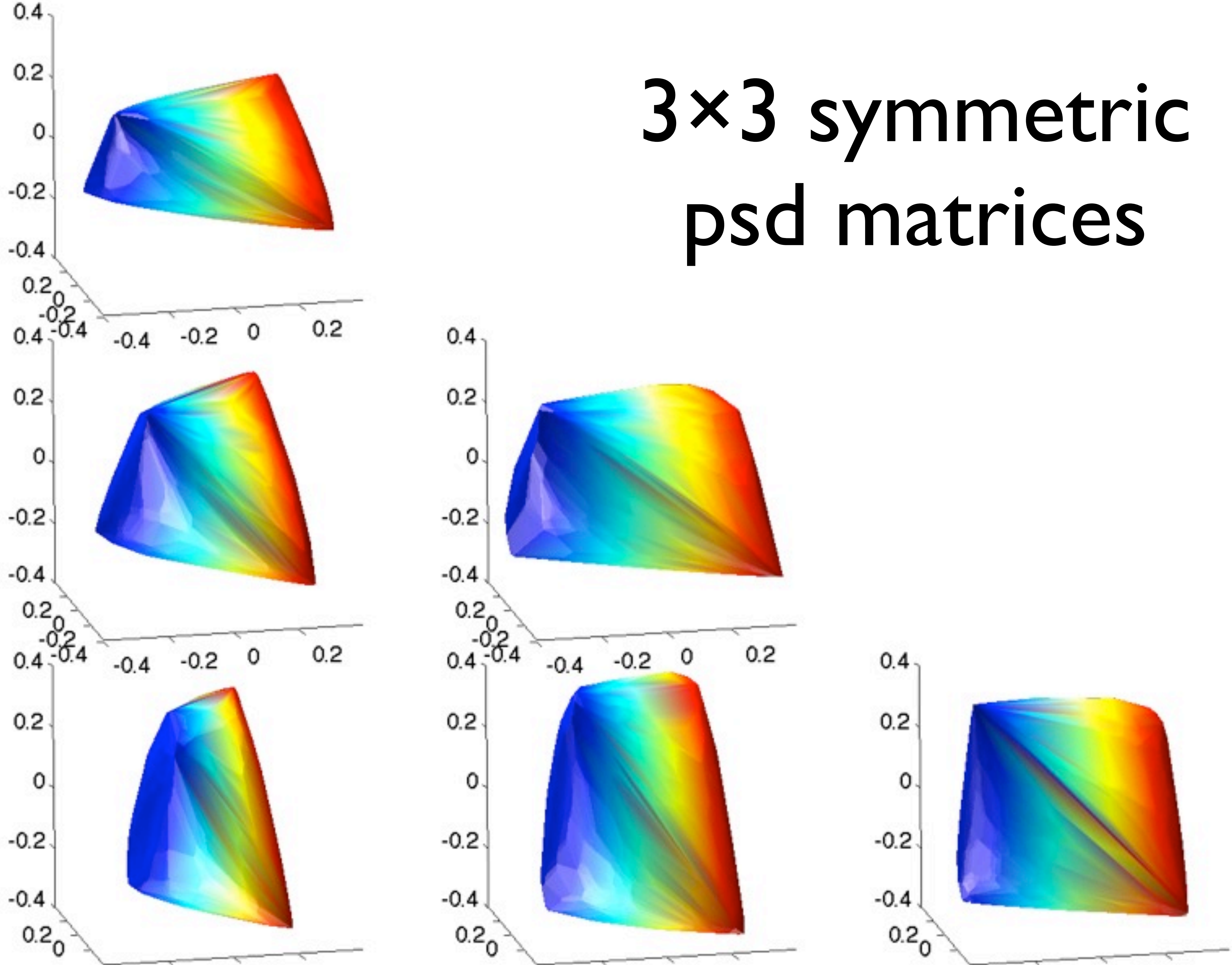
What about 3 x 3?

- Try setting entire diagonal to 1/3
  - plot off-diagonal elements (3 of them)

$$A = \begin{pmatrix} 1/3 & a & b \\ a & 1/3 & c \\ b & c & 1/3 \end{pmatrix}$$



3×3 symmetric psd matrices



S_+ is self-dual

S_+^m

- S_+: { A | $A=A^T$, $x^TAx \geq 0$ for all x }
- [$x^TAx \geq 0$ for all x] $\Leftrightarrow$
  [$tr(B^TA) \geq 0$ for all $\underbrace{\text{psd B}}_{K^*}$]

$A : B \geq 0$

$\Rightarrow$ psd $B = \sum_i x_i x_i^T$ $\quad tr(B^TA) = \sum_i tr((x_i x_i^T)^T A) = \sum_i tr(x_i^T A x_i)$

$= \sum_i x_i^T A x_i \geq 0$

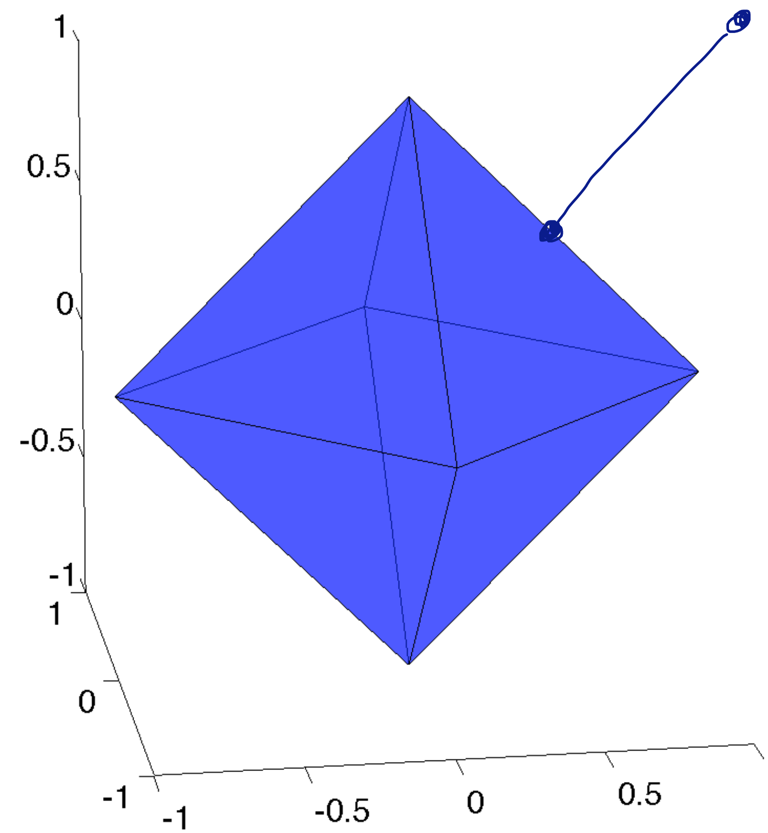
$\Leftarrow$ $B = xx^T$ $\quad tr(B^TA) = x^TAx \geq 0$

How hard are QPs and CPs?

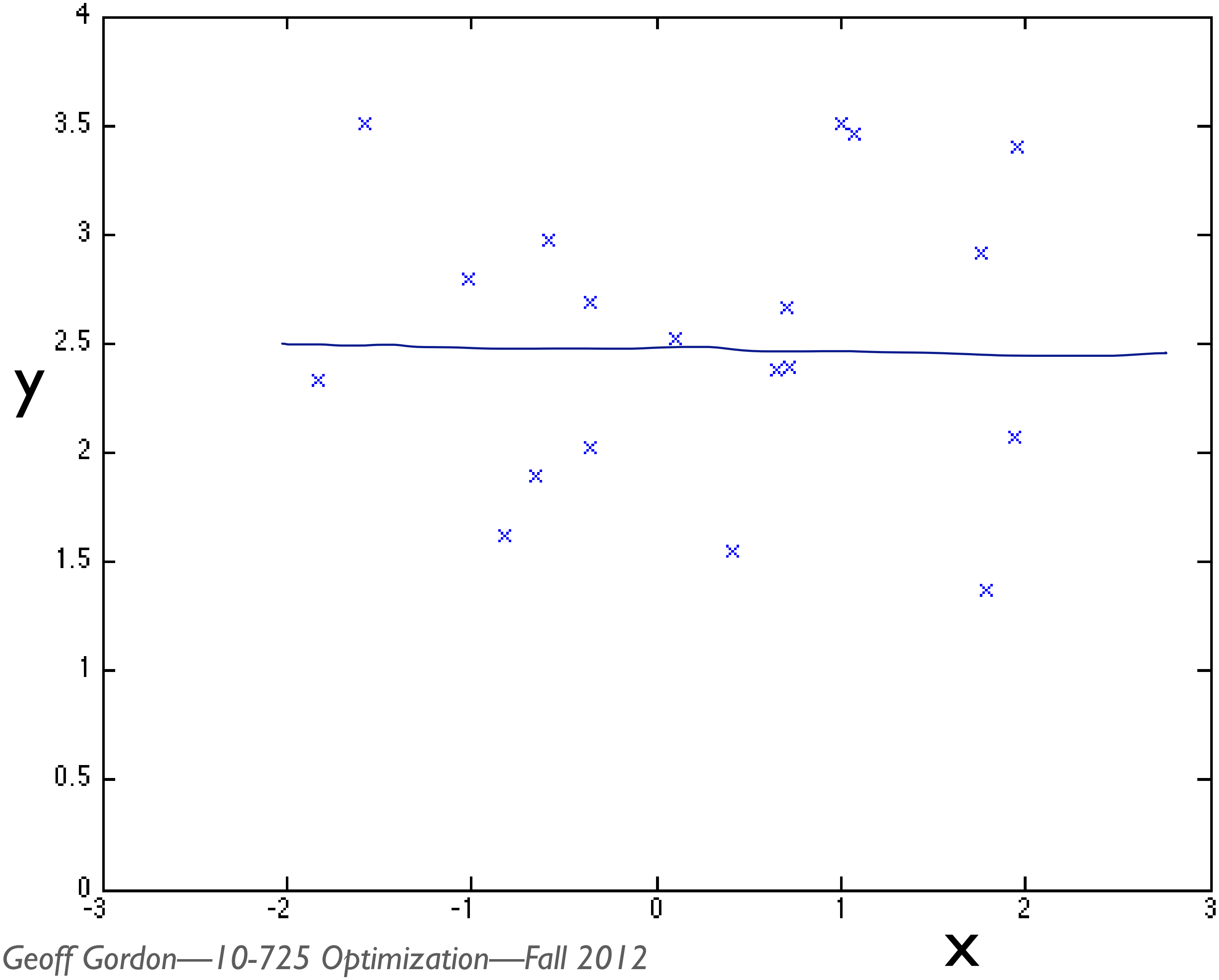
- Convex QP or CP: not much harder than LP!
  - as long as we have an efficient rep'n of the cone
  - poly(L, $1/\epsilon$)   (L = bit length, ϵ = accuracy)
  - can we get strongly polynomial (no $1/\epsilon$)?
    - famous open question, even for LP
- General QP or CP: NP-complete
  - e.g., reduce max cut to QP

QP examples

- Euclidean projection
- LASSO
  - Mahalanobis projection
- Huber regression
- Support vector machine



LASSO example

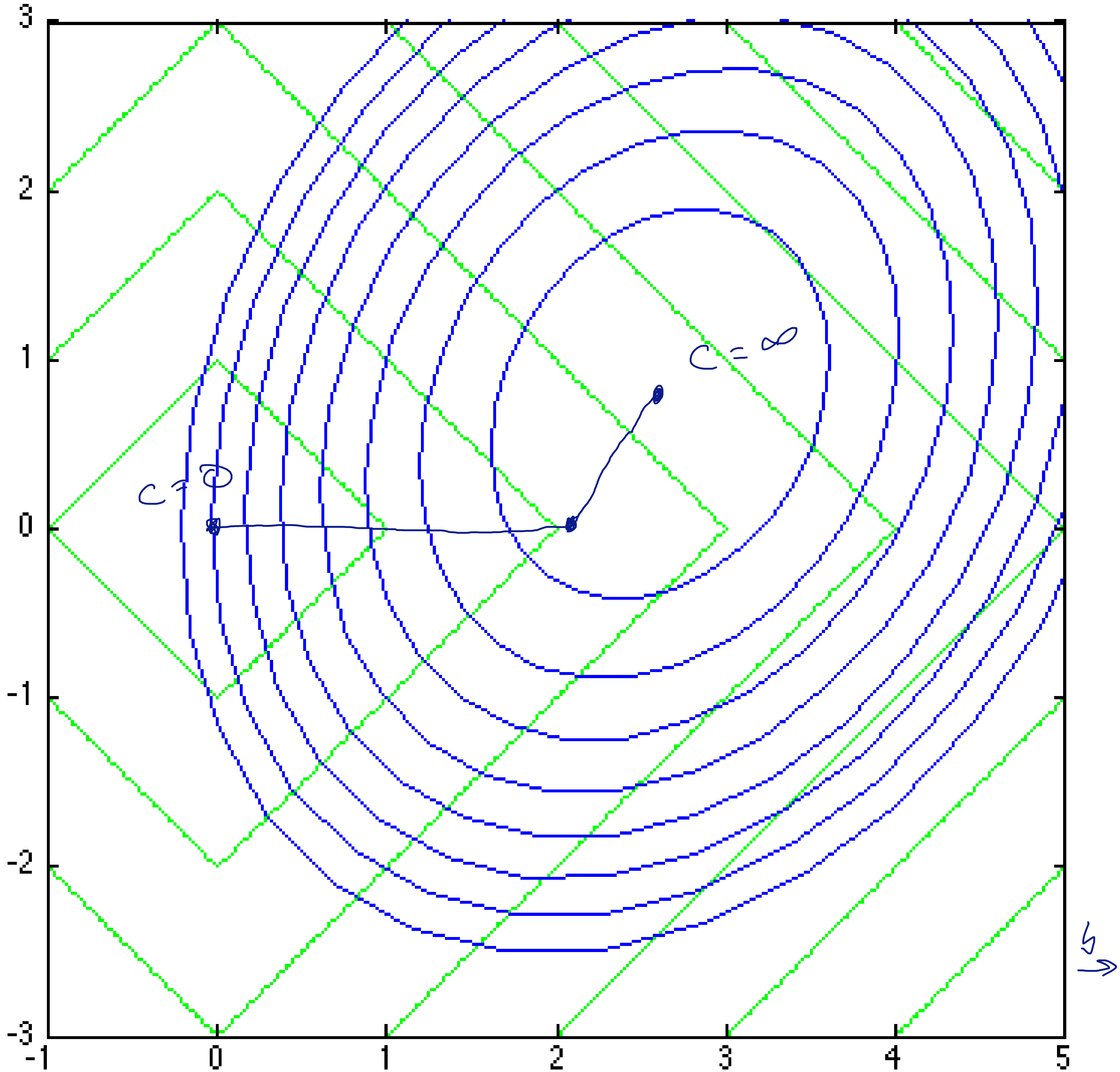


fit y = ax+b
w/ (a,b) sparse

LASSO example

$\|w\|_1 \le C$

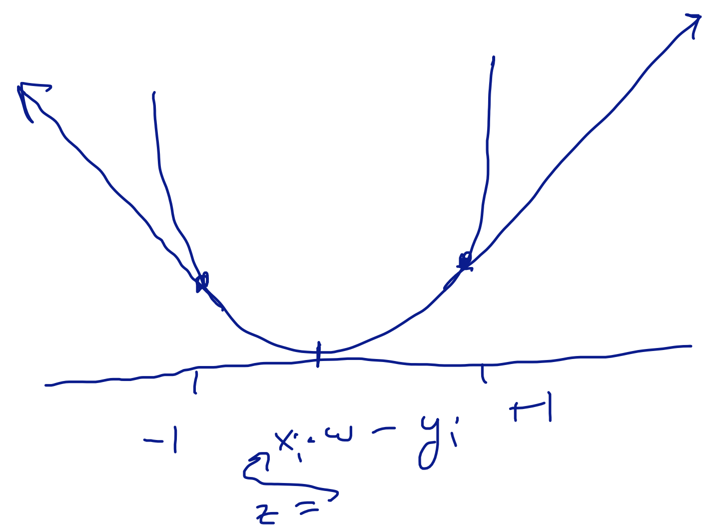
$\uparrow a$



Robust (Huber) regression

- Given points (x_i, y_i) $x_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$
  - L_2 regression: $\min_w \Sigma_i (y_i - x_i^T w)^2$
- Problem: overfitting!   due to outliers
- Solution: Huber loss
  - $\min_w \Sigma_i \mathrm{Hu}(y_i - x_i^T w)$

$$\mathrm{Hu}(z) = \begin{cases} 2z - 1 & z \geq 1 \\ -2z - 1 & z \leq -1 \\ z^2 & -1 \leq z \leq 1 \end{cases}$$



Huber loss as QP

- $\mathrm{Hu}(z_i) = \min_{a,b} (z_i + a_i - b_i)^2 + 2a_i + 2b_i$
  - s.t. $a, b \geq 0$

$a = b = 0 \Rightarrow z^2$

$\nabla_a \quad (z + a - b)^2 + 2a + 2b \quad = \quad 2(z + a - b) + 2 \quad = \quad 0$

$\nabla_b \quad '' \quad = \quad -2(z + a - b) + 2 \quad = \quad 0$

$a > 0 \quad b > 0 \Rightarrow \times$

$a > 0 \quad b = 0 \Rightarrow \quad a = \cancel{b} - z - 1 \qquad z \leq -1$

$(\cancel{z} - \cancel{z} - 1)^2 + 2(-z - 1)$

$-1 - 2z$

Cone program examples

$g_i \subseteq \{1..n\}$

$\min \|y - Xw\|_2^2 + \lambda \sum_i \|w_{g_i}\|_2 \; (+\mu\|w\|_1)$

- SOCP

$\min t + \lambda \sum_i t_i \;\; \text{st} \;\; t \geq \|y - Xw\|_2^2$

$t_i \geq \|w_{g_i}\|_2$

  - (sparse) group lasso
  - discrete MRF relaxation
    - [Kumar, Kolmogorov, Torr, JMLR 2008]
  - min volume covering ellipsoid (nonlinear objective)

$\{x \mid \|Ax + b\| \leq 1\}$ $\quad \begin{pmatrix} Ax_i + b \\ 1 \end{pmatrix} \in \text{SOC}$

$\min \ln|\det A^{-1}|$

Cone program examples

empirical cov

$\min \; \mathrm{tr}(S^T X) - \log\det X + \lambda \sum_{i \neq j} |X_{ij}| \qquad X \succeq 0, \; X = X^T$

- SDP
  - graphical lasso (nonlinear objective)
  - Markowitz portfolio optimization (see B&V)
  - max-cut relaxation [Goemans, Williamson]
  - matrix completion
  - manifold learning: max variance unfolding

Matrix completion

- Observe A_{ij} for $ij \in E$, write $\cancel{P}\,\Omega_{ij} = \begin{cases} 1 & ij \in E \\ 0 & \text{o/w} \end{cases}$

- $\min_X \|(X-A)\circ \cancel{P}\,\Omega\|_F^2 + \underbrace{\lambda\|X\|_*}_{\lambda(\mathrm{tr}(P)+\mathrm{tr}(Q))/2}$

$$M = \begin{bmatrix} P & X \\ X^T & Q \end{bmatrix} \succeq 0 \quad \text{s.t.}$$

$$X = U\Sigma V^T$$

$$M \succeq 0 \iff \mathrm{tr}(B^T M) \geq 0 \;\; \forall B \succeq 0 \qquad \text{take } B = \begin{pmatrix} UU^T & -UV^T \\ -VU^T & VV^T \end{pmatrix}$$

$$0 \leq \mathrm{tr}(B^T M) = \mathrm{tr}(\cancel{UU^T} P) + \mathrm{tr}(\cancel{VV^T} Q) - 2\,\mathrm{tr}(VU^T X)$$

$$\mathrm{tr}(P) + \mathrm{tr}(Q) \geq 2\,\mathrm{tr}(\underbrace{U^T X V}_{\Sigma}) = 2\|X\|_*$$

$$P = U\Sigma U^T \qquad Q = V\Sigma V^T$$

$$\mathrm{tr}(P) = \mathrm{tr}(\cancel{U^T U}\Sigma) \qquad \mathrm{tr}(P) + \mathrm{tr}(Q) = 2\|X\|_*$$

$$\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}^T \begin{pmatrix} U\Sigma U^T & -U\Sigma V^T \\ -V\Sigma U^T & V\Sigma V^T \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} \Sigma & -\Sigma \\ -\Sigma & \Sigma \end{pmatrix} \succeq 0$$