Determinant maximization with linear matrix inequality constraints

S. Boyd, L. Vandenberghe, S.-P. Wu

Information Systems Laboratory
Stanford University

SCCM Seminar 5 February 1996

MAXDET problem definition

$$\begin{array}{ll} \text{minimize} & c^T x + \log\det G(x)^{-1} \\ \text{subject to} & G(x) \triangleq G_0 + x_1 G_1 + \cdots + x_m G_m > 0 \\ & F(x) \triangleq F_0 + x_1 F_1 + \cdots + x_m F_m \geq 0 \end{array}$$

- $x \in \mathbf{R}^m$ is variable
- $G_i = G_i^T \in \mathbf{R}^{l \times l}$, $F_i = F_i^T \in \mathbf{R}^{n \times n}$
- $F(x) \geq 0$, $G(x) > 0$ called *linear matrix inequalities*

- looks specialized, but includes wide variety of convex optimization problems
- convex problem
  - tractable, in theory and practice
  - useful duality theory

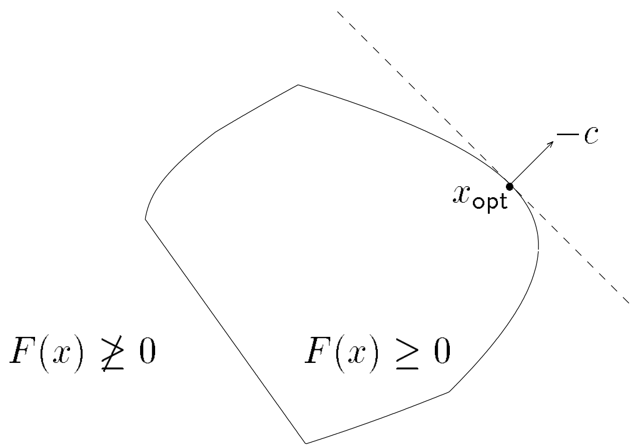
Outline

1. examples of MAXDET probems
2. duality theory
3. interior-point methods

Special cases of MAXDET

semidefinite program (SDP)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m \geq 0 \end{array}$$



LMI can represent many convex constraints
linear inequalities, convex quadratic inequalities, matrix norm constraints, ...

linear program

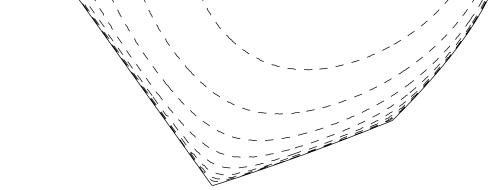
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \ i = 1, \ldots, n \end{array}$$

SDP with $F(x) = \textbf{diag}\,(b - Ax)$

analytic center of LMI

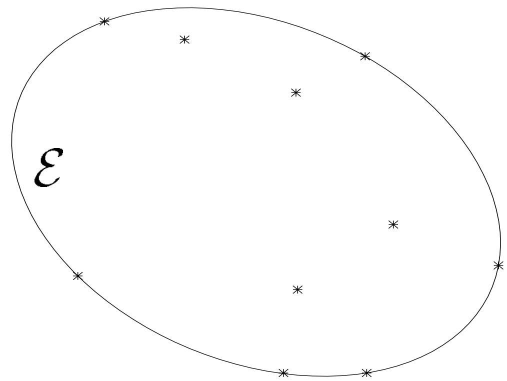
$$\begin{array}{ll} \text{minimize} & \log\det F(x)^{-1} \\ \text{subject to} & F(x) \triangleq F_0 + x_1 F_1 + \cdots + x_m F_m > 0 \end{array}$$

- $\log\det F(x)^{-1}$ smooth, convex on $\{x \mid F(x) > 0\}$
- optimal point x_{ac} maximizes $\det F(x)$
- x_{ac} called analytic center of LMI $F(x) > 0$



Minimum volume ellipsoid around points

find min vol ellipsoid containing points $x_1, \ldots, x_K \in \mathbf{R}^n$



ellipsoid $\mathcal{E} = \{x \mid \|Ax - b\| \leq 1\}$

– center $A^{-1}b$

– $A = A^T > 0$, volume proportional to $\det A^{-1}$

$$\begin{array}{ll} \text{minimize} & \log\det A^{-1} \\ \text{subject to} & A = A^T > 0 \\ & \|Ax_i - b\| \leq 1, \quad i = 1, \ldots, K \end{array}$$

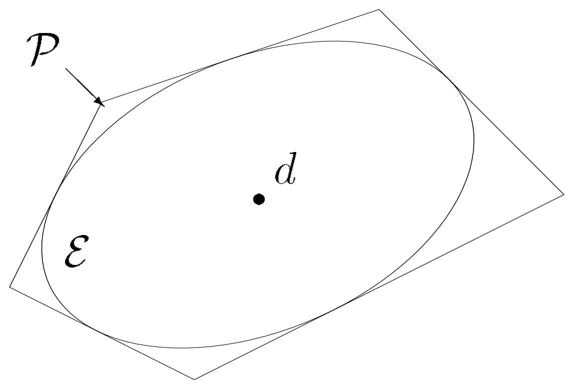
convex optimization problem in A, b
$(n + n(n+1)/2$ vars)

express constraints as LMI

$$\|Ax_i - b\| \leq 1 \Longleftrightarrow \left[\begin{array}{cc} I & Ax_i - b \\ (Ax_i - b)^T & 1 \end{array}\right] \geq 0$$

Maximum volume ellipsoid in polytope

find max vol ellips. in $\mathcal{P} = \{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, L\}$



ellipsoid $\mathcal{E} = \{By + d \mid \|y\| \leq 1\}$

- center d
- $B = B^T > 0$, volume proportional to $\det B$

$$\begin{aligned}
\mathcal{E} \subseteq \mathcal{P} \iff & a_i^T(By + d) \leq b_i \text{ for all } \|y\| \leq 1 \\
\iff & \sup_{\|y\| \leq 1} a_i^T By + a_i^T d \leq b_i \\
\iff & \|Ba_i\| + a_i^T d \leq b_i, \quad i = 1, \ldots, L
\end{aligned}$$

convex constraint in B and d

maximum volume $\mathcal{E} \subseteq \mathcal{P}$

formulation as convex problem in variables B, d:

$$\begin{array}{ll} \text{maximize} & \log\det B \\ \text{subject to} & B = B^T > 0 \\ & \|Ba_i\| + a_i^T d \leq b_i, \ i = 1, \ldots, L \end{array}$$

express constraints as LMI in B, d

$$\|Ba_i\| + a_i^T d \leq b_i \Longleftrightarrow \begin{bmatrix} (b_i - a_i^T d)I & Ba_i \\ (Ba_i)^T & b_i - a_i^T d \end{bmatrix} \geq 0$$

hence, formulation as MAXDET-problem

$$\begin{array}{ll} \text{minimize} & \log\det B^{-1} \\ \text{subject to} & B > 0 \\ & \begin{bmatrix} (b_i - a_i^T d)I & Ba_i \\ (Ba_i)^T & b_i - a_i^T d \end{bmatrix} \geq 0, \quad i = 1, \ldots, L \end{array}$$

Experiment design

estimate x from measurements

$$y_k = a_k^T x + w_k, \quad i = 1, \ldots, N$$

- $a_k \in \{v_1, \ldots, v_m\}$, v_i given test vectors
- w_k IID $N(0,1)$ measurement noise
- $\lambda_i =$ fraction of a_k's equal to v_i
- $N \gg m$

LS estimator: $\widehat{x} = \left(\sum_{k=1}^{N} a_k a_k^T\right)^{-1} \sum_{i=1}^{N} y_k a_k$

error covariance

$$\mathbf{E}(\widehat{x} - x)(\widehat{x} - x)^T = \frac{1}{N}\left(\sum_{i=1}^{m} \lambda_i v_i v_i^T\right)^{-1} = \frac{1}{N} E(\lambda)$$

optimal experiment design: choose λ_i

$$\lambda_i \geq 0, \quad \sum_{i=1}^{m} \lambda_i = 1,$$

that make $E(\lambda)$ 'small'

- minimize $\lambda_{\max}(E(\lambda))$ (E-optimality)
- minimize $\mathbf{Tr}\, E(\lambda)$ (A-optimality)
- minimize $\det E(\lambda)$ (D-optimality)

all are MAXDET problems

D-optimal design

$$\begin{array}{ll}\text{minimize} & \log\det\left(\sum_{i=1}^m \lambda_i v_i v_i^T\right)^{-1} \\ \text{subject to} & \lambda_i \geq 0, \quad i = 1, \ldots, m \\ & \sum_{i=1}^m \lambda_i = 1 \\ & \sum_{i=1}^m \lambda_i v_i v_i^T > 0 \end{array}$$

can add other convex constraints, *e.g.*,

– bounds on cost or time of measurements:

$$c_i^T \lambda \leq b_i$$

– no more than 80% of the measurements is concentrated in less than 20% of the test vectors

$$\sum_{i=1}^{\lfloor m/5 \rfloor} \lambda_{[i]} \leq 0.8$$

($\lambda_{[i]}$ is ith largest component of λ)

Positive definite matrix completion

matrix $A = A^T$

- entries A_{ij}, $(i,j) \in \mathcal{N}$ are fixed
- entries A_{ij}, $(i,j) \notin \mathcal{N}$ are free

positive definite completion

choose free entries such that $A > 0$ (if possible)

maximum entropy completion

$$\begin{array}{ll} \text{maximize} & \log\det A \\ \text{subject to} & A > 0 \end{array}$$

property: $(A^{-1})_{ij} = 0$ for $i, j \notin \mathcal{N}$

(since $\frac{\partial \log\det A^{-1}}{\partial A_{ij}} = -(A^{-1})_{ij}$)

Moment problem

there exists a probability distribution on **R** such that

$$\mu_i = \mathbf{E}t^i, \quad i = 1, \ldots, 2n$$

if and only if

$$H(\mu) = \begin{bmatrix} 1 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_n & \mu_{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} & \mu_{2n} \end{bmatrix} \geq 0$$

LMI in variables μ_i

hence, can solve

$$\begin{array}{ll} \text{maximize/minimize} & \mathbf{E}(c_0 + c_1 t + \cdots + c_{2n} t^{2n}) \\ \text{subject to} & \underline{\mu}_i \leq \mathbf{E}t^i \leq \overline{\mu}_i, \quad i = 1, \ldots, 2n \end{array}$$

over all probability distributions on **R** by solving SDP

$$\begin{array}{ll} \text{maximize/minimize} & c_0 + c_1\mu_1 + \cdots + c_{2n}\mu_{2n} \\ \text{subject to} & \underline{\mu}_i \leq \mu_i \leq \overline{\mu}_i, \quad i = 1, \ldots, 2n \\ & H(\mu_1, \ldots, \mu_{2n}) \geq 0 \end{array}$$

Other applications

- maximizing products of positive concave functions
- minimum volume ellipsoid covering union or sum of ellipsoids
- maximum volume ellipsoid in intersection or sum of ellipsoids
- computing channel capacity in information theory
- maximum likelihood estimation

MAXDET duality theory

primal MAXDET problem

$$\begin{array}{ll} \text{minimize} & c^T x + \log\det G(x)^{-1} \\ \text{subject to} & G(x) = G_0 + x_1 G_1 + \cdots + x_m G_m > 0 \\ & F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m \geq 0 \end{array}$$

optimal value $p^\star$

dual MAXDET problem

$$\begin{array}{ll} \text{maximize} & \log\det W - \mathbf{Tr}\, G_0 W - \mathbf{Tr}\, F_0 Z + l \\ \text{subject to} & \mathbf{Tr}\, F_i Z + \mathbf{Tr}\, G_i W = c_i, \quad i = 1, \ldots, m \\ & W > 0, \quad Z \geq 0 \end{array}$$

variables $W = W^T \in \mathbf{R}^{l \times l}$, $Z = Z^T \in \mathbf{R}^{n \times n}$
optimal value $d^\star$

properties

- $p^\star \geq d^\star$ (always)
- $p^\star = d^\star$ (usually)

definition

duality gap $=$ primal objective $-$ dual objective

Example: experiment design

primal problem

$$\begin{array}{ll} \text{minimize} & \log\det\left(\sum_{i=1}^m \lambda_i v_i v_i^T\right)^{-1} \\ \text{subject to} & \sum_{i=1}^m \lambda_i = 1 \\ & \lambda_i \geq 0, \quad i = 1, \ldots, m \\ & \sum_{i=1}^m \lambda_i v_i v_i^T > 0 \end{array}$$

dual problem

$$\begin{array}{ll} \text{maximize} & \log\det W \\ \text{subject to} & W = W^T > 0 \\ & v_i^T W v_i \leq 1, \quad i = 1, \ldots, m \end{array}$$

interpretation: W determines smallest ellipsoid with center at the origin and containing v_i, $i = 1, \ldots, m$

Central path: general

general convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in C \end{array}$$

f_0, C convex

φ is **barrier function** for C

– smooth, convex

– $\varphi(x) \to \infty$ as $x(\in \textbf{int}\, C) \to \partial C$

central path

$$x^\star(t) = \mathop{\textbf{argmin}}_{x \in C} \left(t f_0(x) + \varphi(x) \right) \quad \text{for } t > 0$$

Central path: MAXDET problem

$$\begin{aligned} f_0(x) &= c^T x + \log\det G(x)^{-1} \\ C &= \{x \mid F(x) \geq 0\} \end{aligned}$$

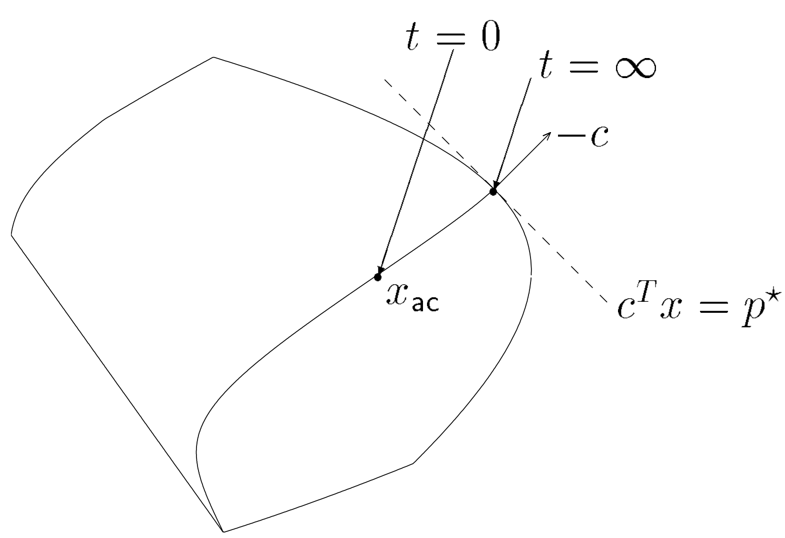
barrier function for LMI $F(x) \geq 0$

$$\varphi(x) = \begin{cases} \log\det F(x)^{-1} & \text{if } F(x) > 0 \\ +\infty & \text{otherwise} \end{cases}$$

MAXDET central path: $x^\star(t) = \mathop{\textbf{argmin}}_{\substack{F(x)>0 \\ G(x)>0}} \varphi(t,x)$, with

$$\varphi(t,x) = t\left(c^T x + \log\det G(x)^{-1}\right) + \log\det F(x)^{-1}$$

example: SDP



Path-following for MAXDET

properties of MAXDET central path

- from $x^\star(t)$, get dual feasible $Z^\star(t)$, $W^\star(t)$
- corresponding duality gap is n/t
- $x^\star(t) \to$ optimal as $t \to \infty$

path-following algorithm

> **given** strictly feasible x, $t \geq 1$
>
> **repeat**
>
> 1. compute $x^\star(t)$ using Newton's method
> 2. $x := x^\star(t)$
> 3. increase t
>
> **until** $n/t <$ tol

tradeoff: large increase in t means

- fast gap reduction (fewer outer iterations), but
- many Newton steps to compute $x^\star(t^+)$
  (more Newton steps per outer iteration)

Complexity of Newton's method

(Nesterov & Nemirovsky, late 1980s)

for **self-concordant** functions

definition: along a line

$$|f'''(t)| \leq K f''(t)^{3/2}$$

Example: $(K = 2)$

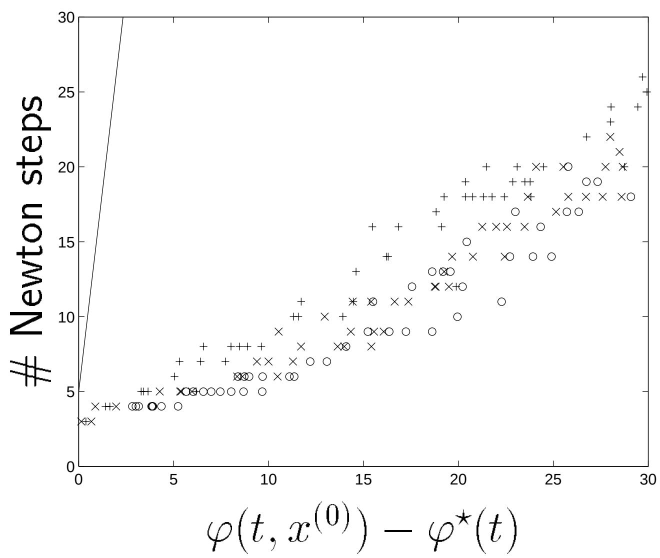
$$\varphi(t, x) = t(c^T x + \log\det G(x)^{-1}) + \log\det F(x)^{-1} \quad (t \geq 1)$$

complexity of Newton's method

- **theorem:** #Newton steps to minimize $\varphi(t, x)$, starting from $x^{(0)}$:

$$\#\text{steps} \leq 10.7(\varphi(t, x^{(0)}) - \varphi^\star(t)) + 5$$

- **empirically**: #steps $\approx (\varphi(t, x^{(0)}) - \varphi^\star(t)) + 3$



Path-following algorithm

idea: choose t^+, starting point $\widehat{x}$ for Newton alg. s.t.

$$\varphi(t^+, \widehat{x}) - \varphi^\star(t^+) = \gamma$$

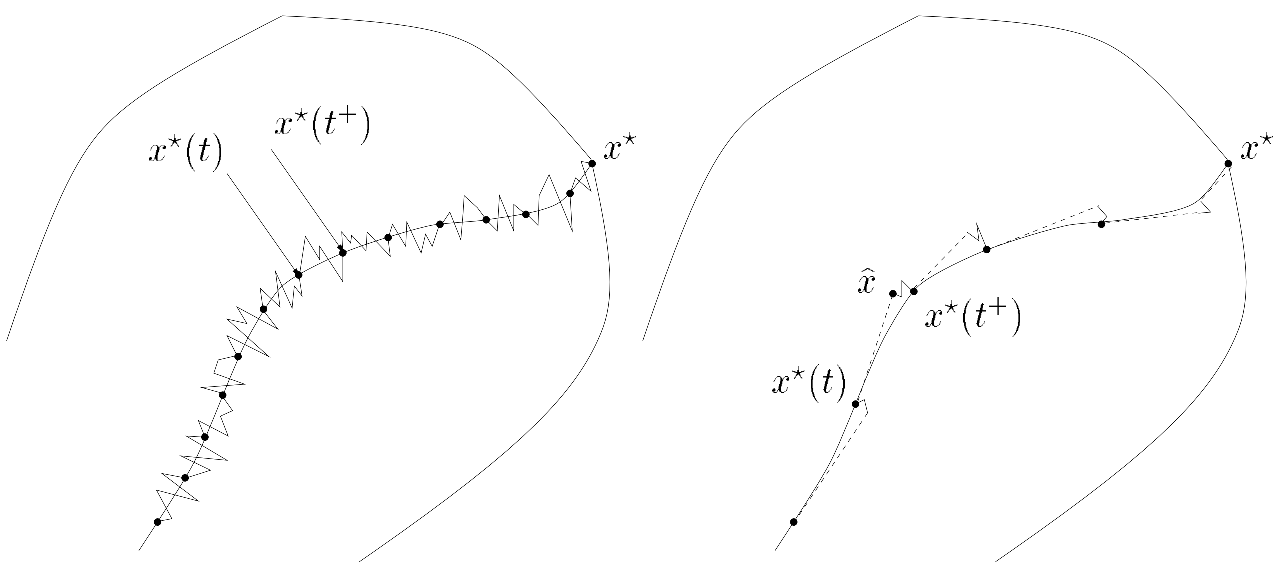
(bounds # Newton steps required to compute $x^\star(t^+)$)

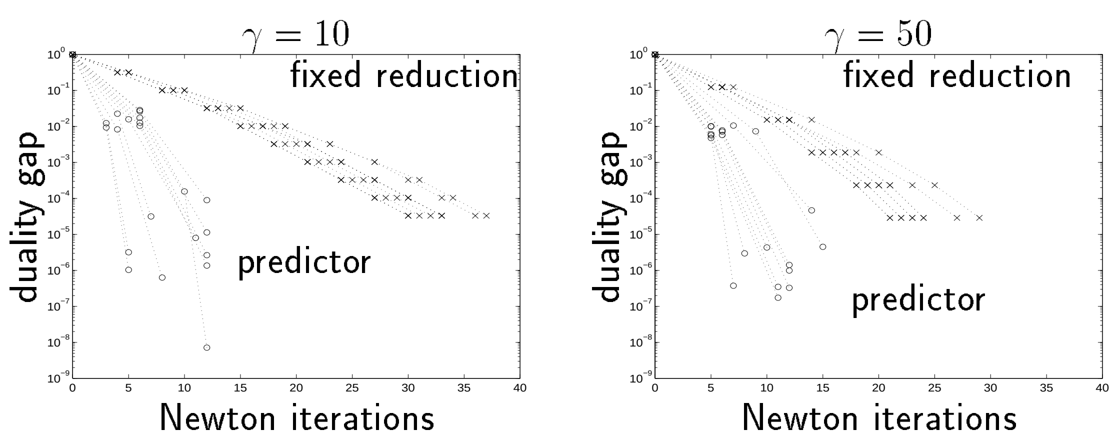
in **practice**: use lower bound from duality

$$\begin{aligned}\varphi(t^+, \widehat{x}) - \varphi^\star(t^+) \leq\ & \varphi(t^+, \widehat{x}) + \log\det Z^{-1} \\ & + t\left(\log\det W^{-1} + \mathbf{Tr}\, G_0 W + \mathbf{Tr}\, F_0 Z - l\right) \\ =\ & \varphi(t^+, \widehat{x}) + \text{function of } W, Z\end{aligned}$$

two extreme choices

- fixed reduction: $\widehat{x} = x^\star(t)$, $t^+ = \left(1 + \sqrt{2\gamma/n}\right) t$
- predictor step along tangent of central path

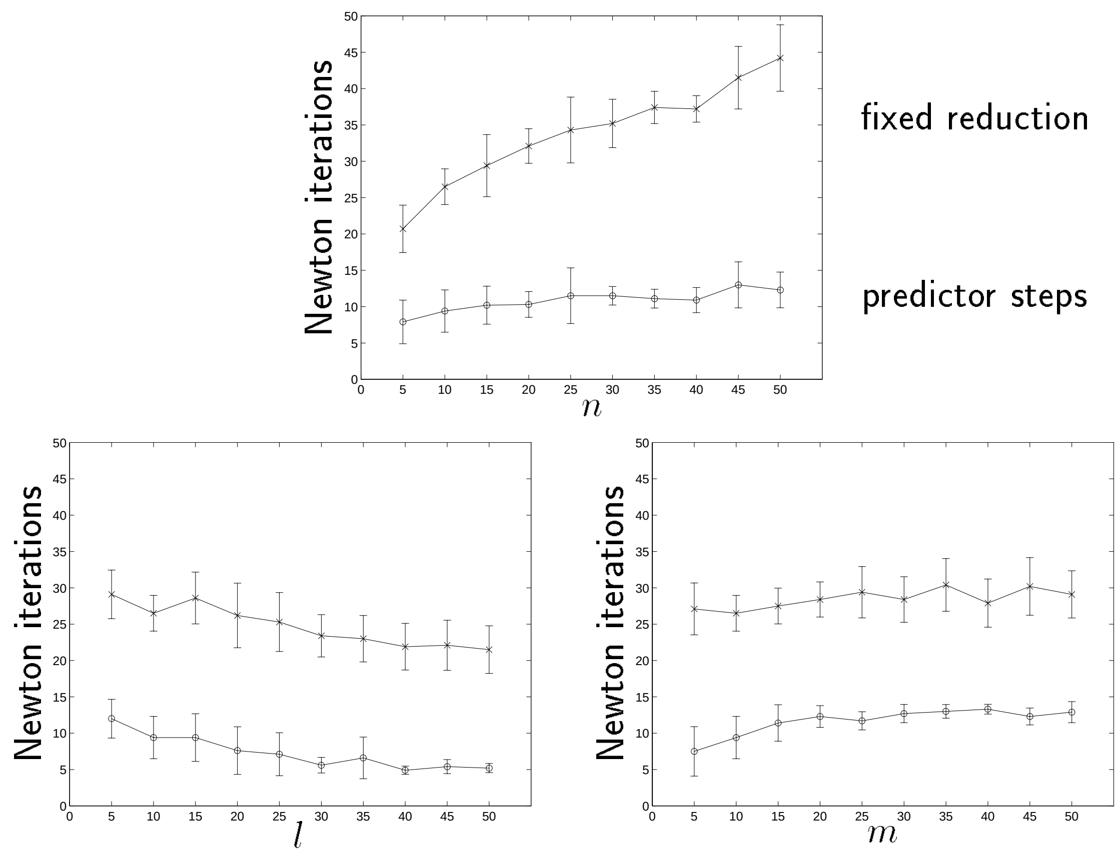




Total complexity

total number of Newton steps

- upper bound: $O\left(\sqrt{n}\log(1/\epsilon)\right)$
- practice, fixed-reduction method: $O\left(\sqrt{n}\log(1/\epsilon)\right)$
- practice, with predictor steps: $O\left(\log(1/\epsilon)\right)$



one Newton step involves a least-squares problem

$$\text{minimize} \quad \left\|\tilde{F}(v)\right\|_F^2 + \left\|\tilde{G}(v)\right\|_F^2$$

Conclusion

MAXDET-problem

$$\begin{array}{ll} \text{minimize} & c^T x + \log\det G(x)^{-1} \\ \text{subject to} & G(x) > 0,\ F(x) \geq 0 \end{array}$$

arises in many different areas

- includes SDP, LP, convex QCQP
- geometrical problems involving ellipsoids
- experiment design, max. likelihood estimation, channel capacity, ...

convex, hence can be solved very efficiently

software/paper available on ftp soon (anonymous ftp to `isl.stanford.edu` in `/pub/boyd/maxdet`)