Engineering Applications of Convex Optimization

Stephen Boyd Michael Grant Almir Mutapcic Kwangmoo Koh

Electrical Engineering Department, Stanford University

Outline

- bounding portfolio risk
- computing probability bounds
- antenna array beamforming
- l_1-regularized logistic regression

Portfolio risk bounding

- portfolio of n assets invested for single period
- w_i is amount of investment in asset i
- returns of assets is random vector r with mean $\overline{r}$, covariance Σ
- portfolio return is random variable $r^T w$
- mean portfolio return is $\overline{r}^T w$; variance is $V = w^T \Sigma w$

value at risk & probability of loss are related to portfolio variance V

Risk bound with uncertain covariance

now suppose:

- w is known (and fixed)
- have only partial information about Σ, $i.e.$,

$$L_{ij} \leq \Sigma_{ij} \leq U_{ij}, \quad i, j = 1, \ldots, n$$

problem: how large can portfolio variance $V = w^T \Sigma w$ be?

Risk bound via semidefinite programming

can get (tight) bound on V via semidefinite programming (SDP):

$$\begin{array}{ll} \text{maximize} & w^T \Sigma w \\ \text{subject to} & \Sigma \succeq 0 \\ & L_{ij} \leq \Sigma_{ij} \leq U_{ij} \end{array}$$

variable is matrix $\Sigma = \Sigma^T$; $\Sigma \succeq 0$ means Σ is positive semidefinite

many extensions possible, $e.g.$, optimize portfolio w with worst-case variance limit

cvx code

```
cvx_begin
    variable Sigma(n,n) symmetric
    maximize ( w'*Sigma*w )
    subject to
    Sigma == semidefinite(n);    % Sigma is positive semidefinite
    Sigma >= L;
    Sigma <= U;
cvx_end
```

Example

portfolio with $n = 4$ assets

variance bounding with sign constraints on Σ:

$$w = \begin{bmatrix} 1 \\ 2 \\ -.5 \\ .5 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 1 & + & + & ? \\ + & 1 & - & - \\ + & - & 1 & + \\ ? & - & + & 1 \end{bmatrix}$$

($i.e.$, $\Sigma_{12} \geq 0$, $\Sigma_{23} \leq 0$, . . .)

Result

(global) maximum value of V is 10.1, with

$$\Sigma = \begin{bmatrix} 1.00 & 0.79 & 0.00 & 0.53 \\ 0.79 & 1.00 & -.59 & 0.00 \\ 0.00 & -.59 & 1.00 & 0.51 \\ 0.53 & 0.00 & 0.51 & 1.00 \end{bmatrix}$$

(which has rank 3, so constraint $\Sigma \succeq 0$ is active)

- $\Sigma = I$ yields $V = 5.5$
- $\Sigma = [(L+U)/2]_+$ yields $V = 6.75$ ($[\cdot]_+$ is positive semidefinite part)

Computing probability bounds

random variable $(X, Y) \in \mathbf{R}^2$ with

- $\mathcal{N}(0, 1)$ marginal distributions
- X, Y uncorrelated

question: how large (small) can $\mathbf{Prob}(X \leq 0, Y \leq 0)$ be?

if $(X, Y) \sim \mathcal{N}(0, I)$, $\mathbf{Prob}(X \leq 0, Y \leq 0) = 0.25$

Probability bounds via LP

- discretize distribution as p_{ij}, $i,j = 1,\ldots,n$, over region $[-3,3]^2$
- $x_i = y_i = 6(i-1)/(n-1) - 3$, $i = 1,\ldots,n$

$$
\begin{array}{ll}
\text{maximize (minimize)} & \sum_{i,j=1}^{n/2} p_{ij} \\
\text{subject to} & p_{ij} \geq 0, \quad i,j = 1,\ldots,n \\
& \sum_{i=1}^{n} p_{ij} = a e^{-y_i^2/2}, \quad j = 1,\ldots,n \\
& \sum_{j=1}^{n} p_{ij} = a e^{-x_i^2/2}, \quad i = 1,\ldots,n \\
& \sum_{i,j=1}^{n} p_{ij} x_i y_j = 0
\end{array}
$$

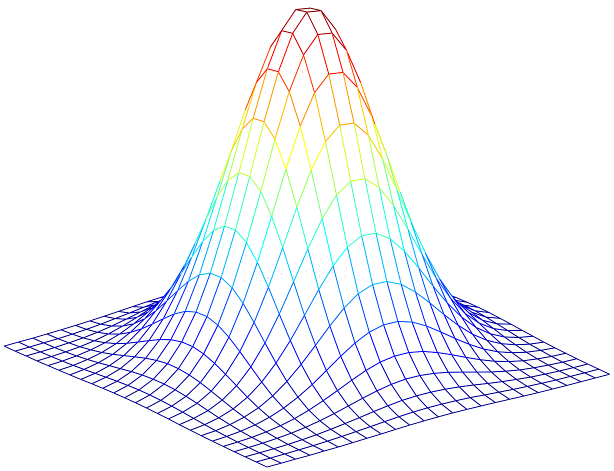
with variable $p \in \mathbf{R}^{n\times n}$, $a = 2.39/(n-1)$

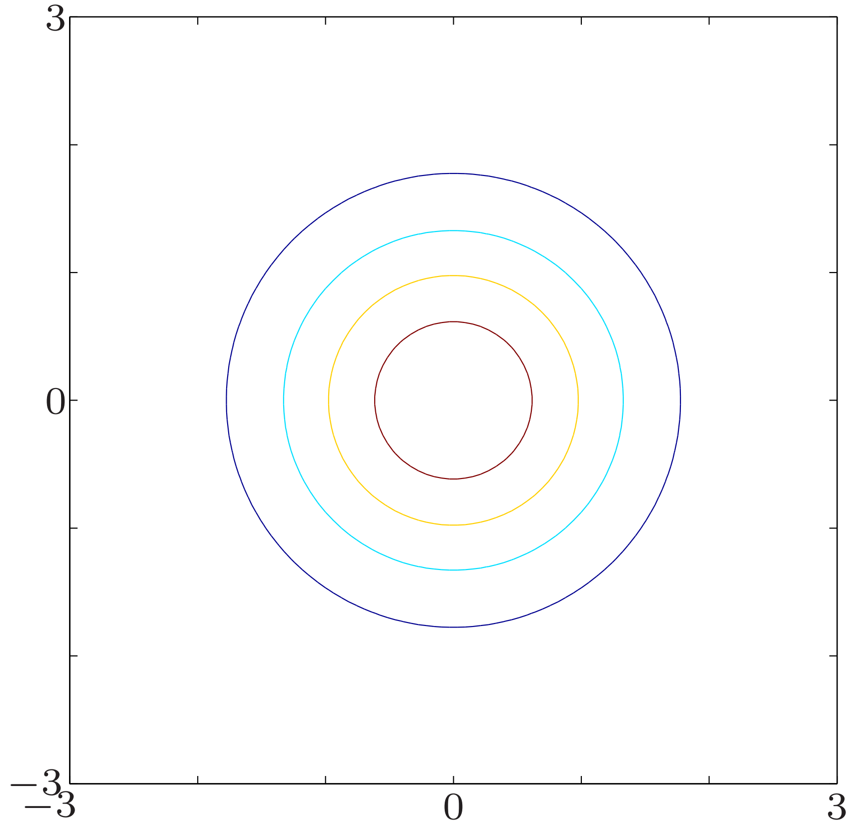
cvx code

```
cvx_begin
    variable p(n,n)
    maximize ( sum(sum(p(1:n/2,1:n/2))) )
    subject to
        p >= 0;
        sum( p,1 ) == a*exp(-y.^2/2)';
        sum( p,2 ) == a*exp(-x.^2/2)';
        sum( sum( p.*(x*y') ) ) == 0;
cvx_end
```

Gaussian

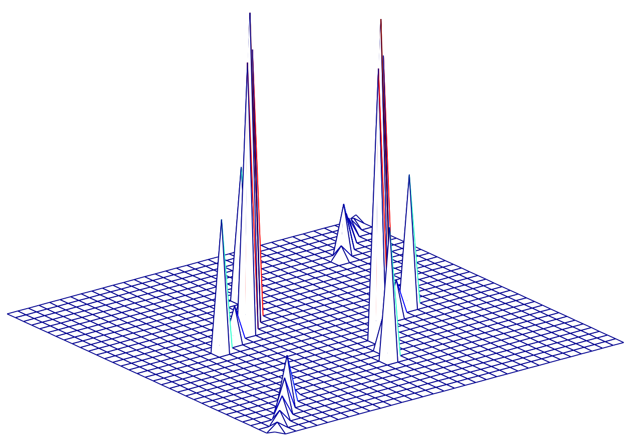
$(X, Y) \sim \mathcal{N}(0.I)$; $\mathbf{Prob}(X \leq 0, Y \leq 0) = 0.25$

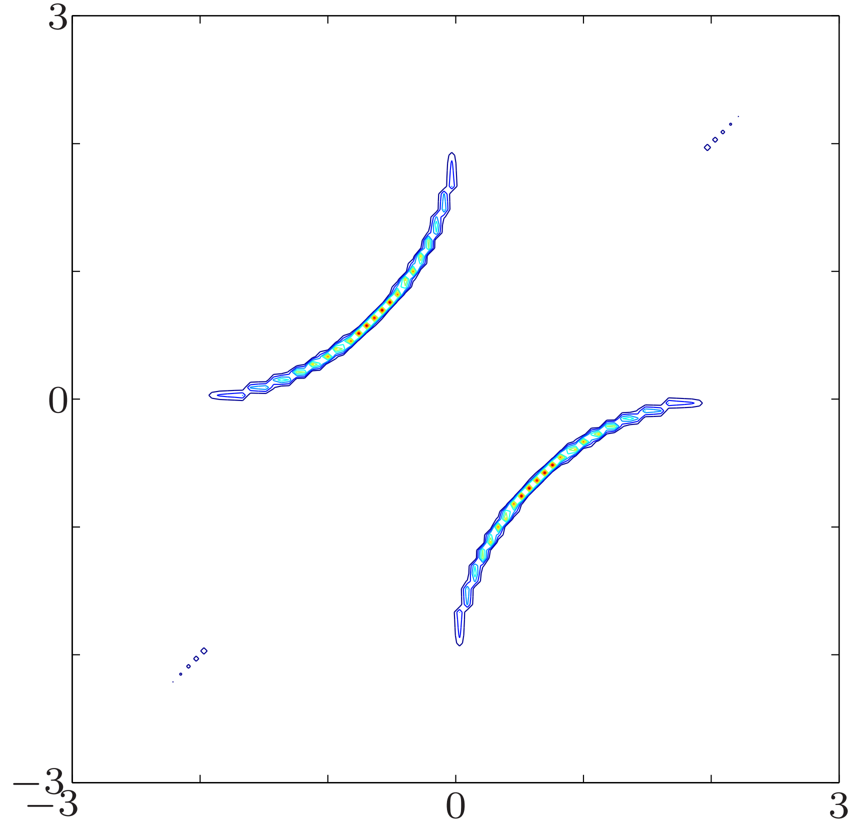




Distribution that minimizes $\mathbf{Prob}(X \leq 0, Y \leq 0)$

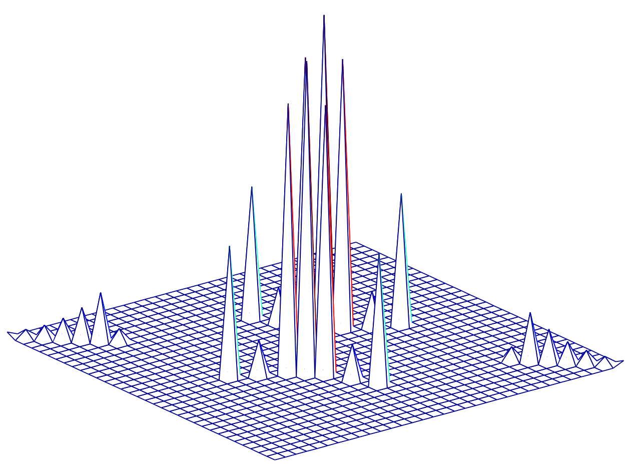
$\mathbf{Prob}(X \leq 0, Y \leq 0) = 0.03$

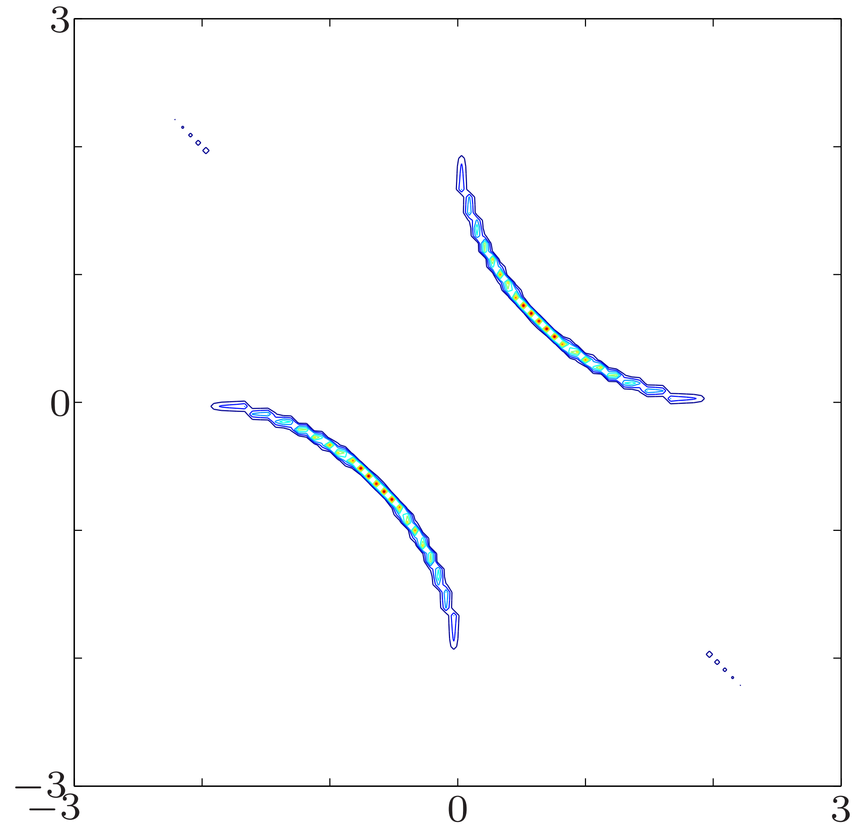




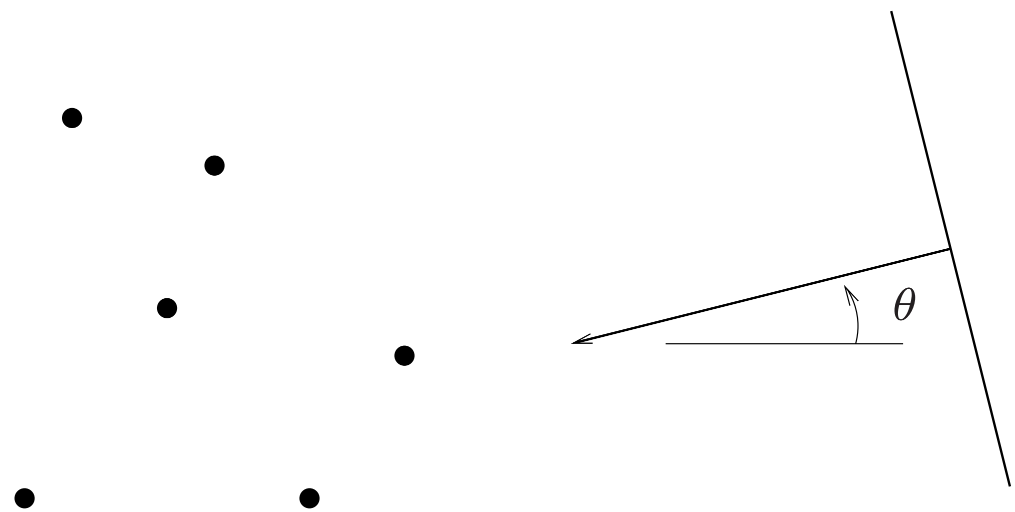
Distribution that maximizes $\mathbf{Prob}(X \leq 0, Y \leq 0)$

$\mathbf{Prob}(X \leq 0, Y \leq 0) = 0.47$





Antenna array beamforming



- n omnidirectional antenna elements in plane, at positions (x_i, y_i)
- unit plane wave ($\lambda = 1$) incident from angle θ
- ith element has (demodulated) signal $e^{j(x_i \cos\theta + y_i \sin\theta)}$ ($j = \sqrt{-1}$)

- combine antenna element signals using complex weights w_i to get antenna array output

$$y(\theta) = \sum_{i=1}^{n} w_i e^{j(x_i \cos\theta + y_i \sin\theta)}$$

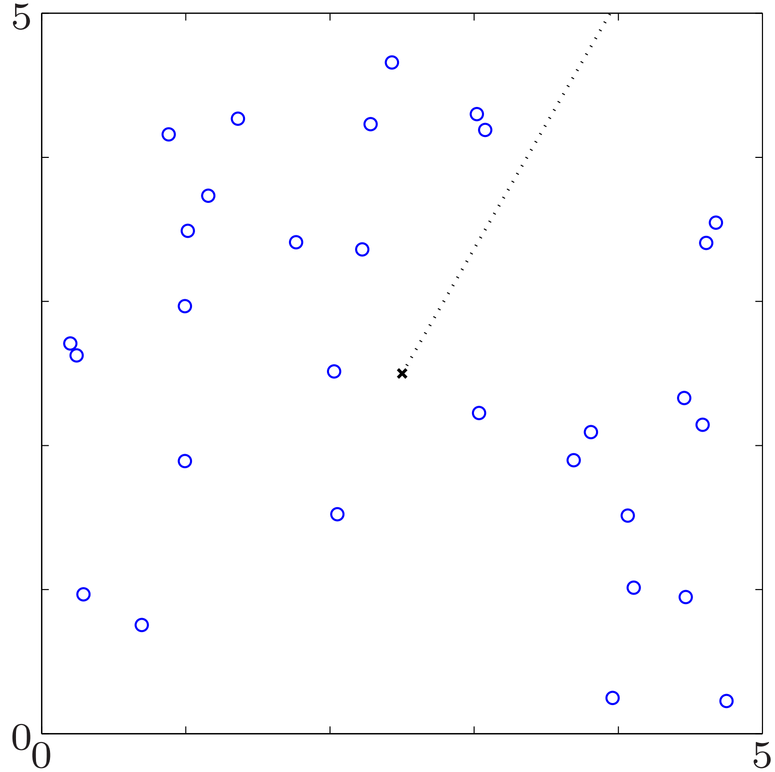
typical design problem:

choose $w \in \mathbf{C}^n$ so that

- $y(\theta_{\rm tar}) = 1$ (unit gain in target or look direction)
- $|y(\theta)|$ is small for $|\theta - \theta_{\rm tar}| \geq \Delta$ (2Δ is beamwidth)

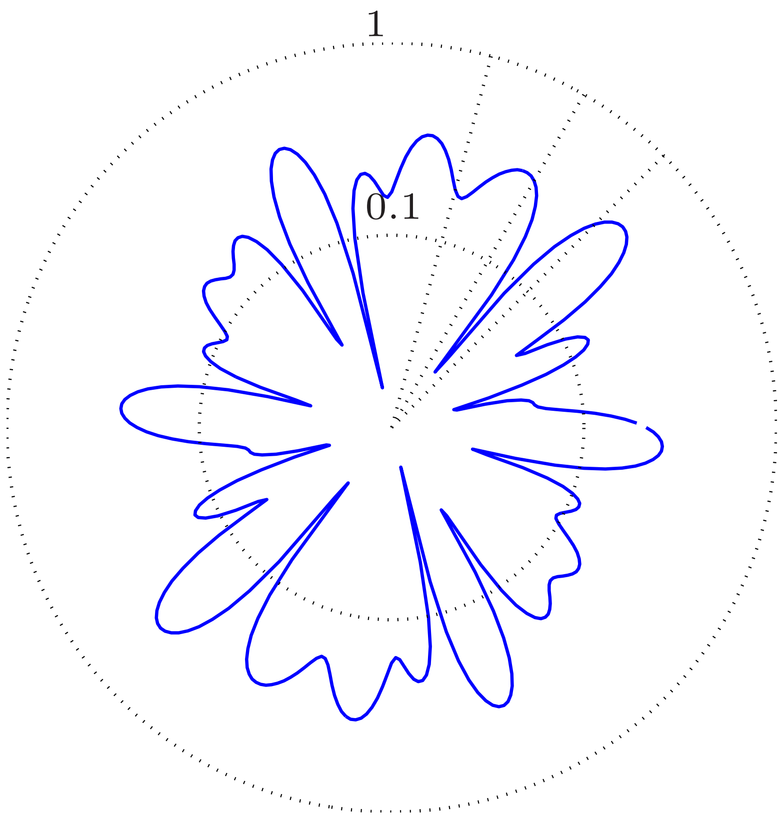
Example

$n = 30$ antenna elements, $\theta_{\mathrm{tar}} = 60^\circ$, $\Delta = 15^\circ$ (30° beamwidth)



Uniform weights

$w_i = 1/n$; no particular directivity pattern



Least-squares beamforming

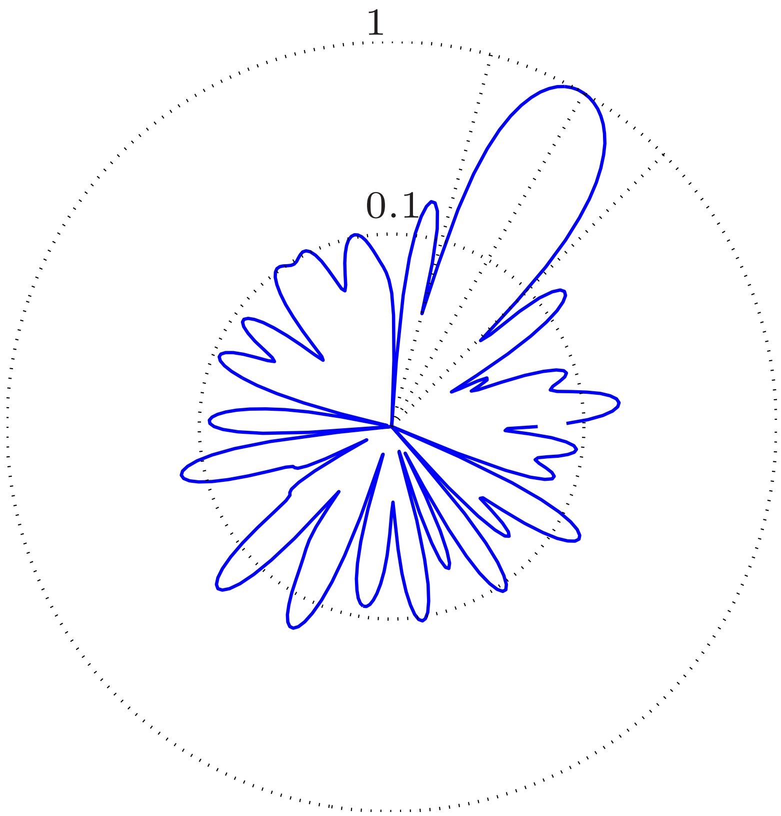
discretize angles outside beam (*i.e.*, $|\theta - \theta_{\rm tar}| \geq \Delta$) as $\theta_1, \ldots, \theta_N$;

solve least-squares problem

$$\begin{array}{ll} \mbox{minimize} & \sum_{i=1}^N |y(\theta_i)|^2 \\ \mbox{subject to} & y(\theta_{\rm tar}) = 1 \end{array}$$

```
cvx_begin
    variable w(n) complex
    minimize ( norm( A_outside_beam*w ) )
    subject to
      a_tar'*w == 1;
cvx_end
```

Least-squares beamforming



Chebyshev beamforming

solve minimax problem

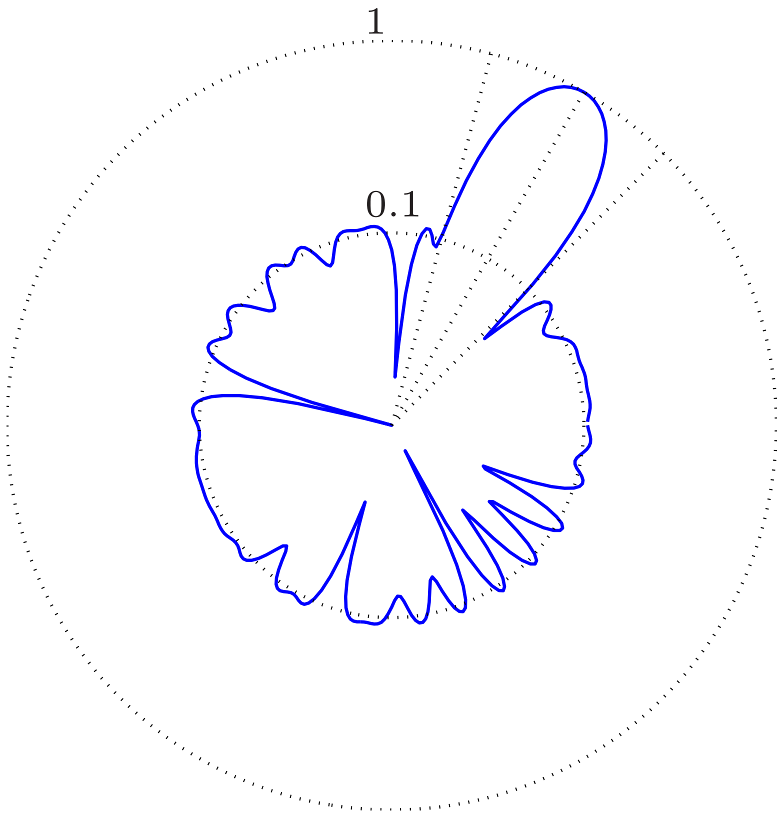
$$\begin{array}{ll} \text{minimize} & \max_{i=1,\ldots,N} |y(\theta_i)| \\ \text{subject to} & y(\theta_{\text{tar}}) = 1 \end{array}$$

(objective is called sidelobe level)

```
cvx_begin
    variable w(n) complex
    minimize ( max( abs( A_outside_beam*w ) ) )
    subject to
      a_tar'*w == 1;
cvx_end
```

Chebyshev beamforming

(globally optimal) sidelobe level 0.11



ℓ_1-regularized logistic regression

logistic model:

$$\mathbf{Prob}(y = 1) = \frac{\exp(a^T x + b)}{1 + \exp(a^T x + b)}$$

- $y \in \{-1, 1\}$ is Boolean random variable (outcome)
- $x \in \mathbf{R}^n$ is vector of explanatory variables or features
- $a \in \mathbf{R}^n$, b are model parameters
- $a^T x + b = 0$ is neutral hyperplane
- linear classifier: given x, $\hat{y} = \mathrm{sgn}(a^T x + b)$

Maximum likelihood estimation

a.k.a. logistic regression

given observed (training) examples $(x_1, y_1) \ldots, (x_m, y_m)$, estimate a, b

maximum likelihood model parameters found by solving (convex) problem

$$\text{minimize} \quad \sum_{i=1}^{n} \log\left(1 + \exp\left(-y_i(x_i^T a + b)\right)\right)$$

with variables $a \in \mathbf{R}^n$, $b \in \mathbf{R}$

ℓ_1-regularized logistic regression

find a, b by solving (convex) problem

$$\text{minimize} \quad \sum_{i=1}^{n} \log\left(1 + \exp\left(-y_i(x_i^T a + b)\right)\right) + \lambda \|a\|_1$$

$\lambda > 0$ is regularization parameter

- protects against over-fitting
- heuristic to get sparse a ($i.e.$, simple explanation) for $m > n$
- heuristic to select relevant features when $m < n$

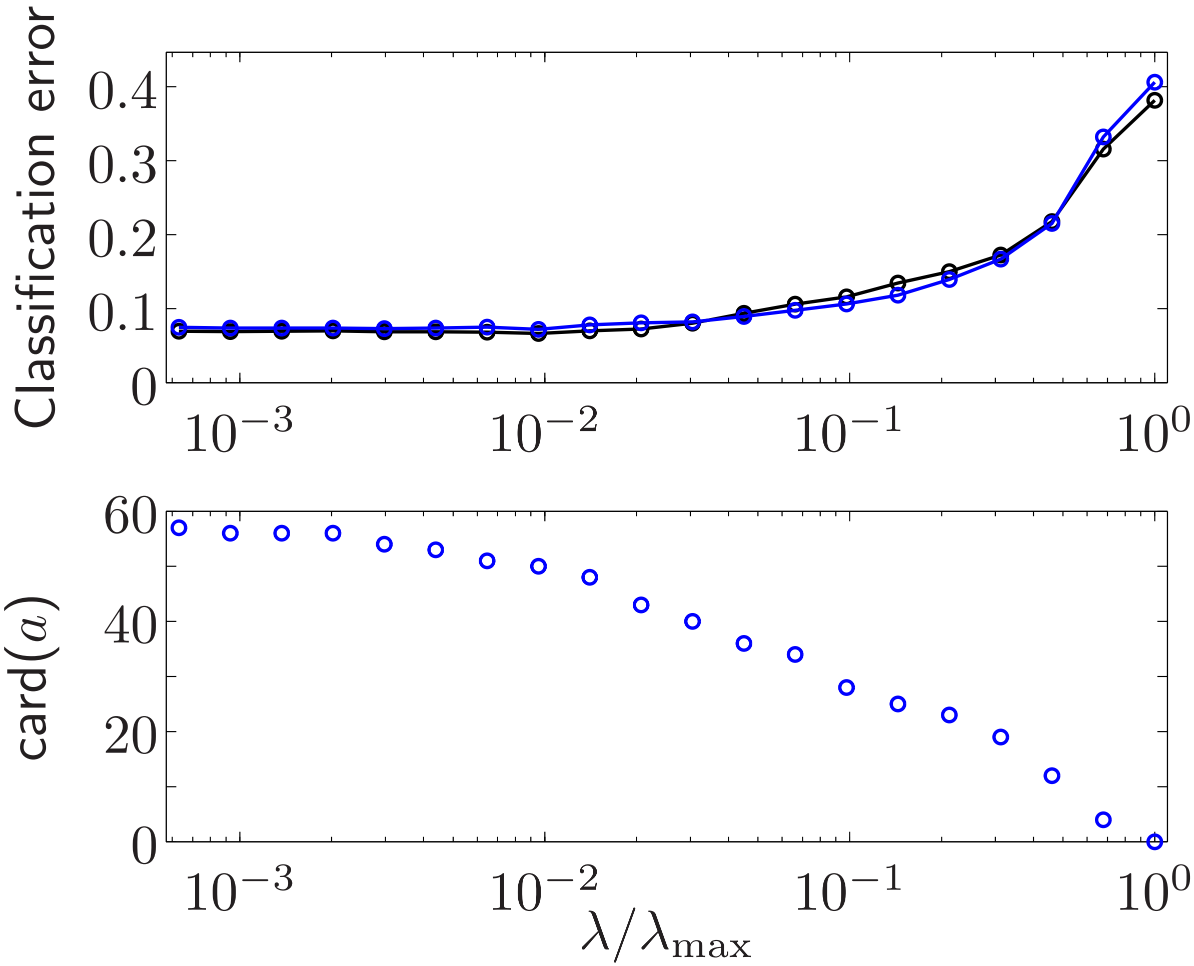
cvx code

```
cvx_begin
    variables a(n) b
    minimize ( sum(log(1+exp(-y.*(X'*a+b))+lambda*norm(a,1) )
cvx_end
```

Spambase example

- taken from UC Irvine machine learning repository
- $n = 57$ features extracted from $m = 4601$ email messages (1813 spam, 2788 nonspam)
- messages randomly divided into training set (2300 examples), validation set (2301 examples)
- ℓ_1-regularized logistic regression model found using training set; classification performance checked on validation set

Results

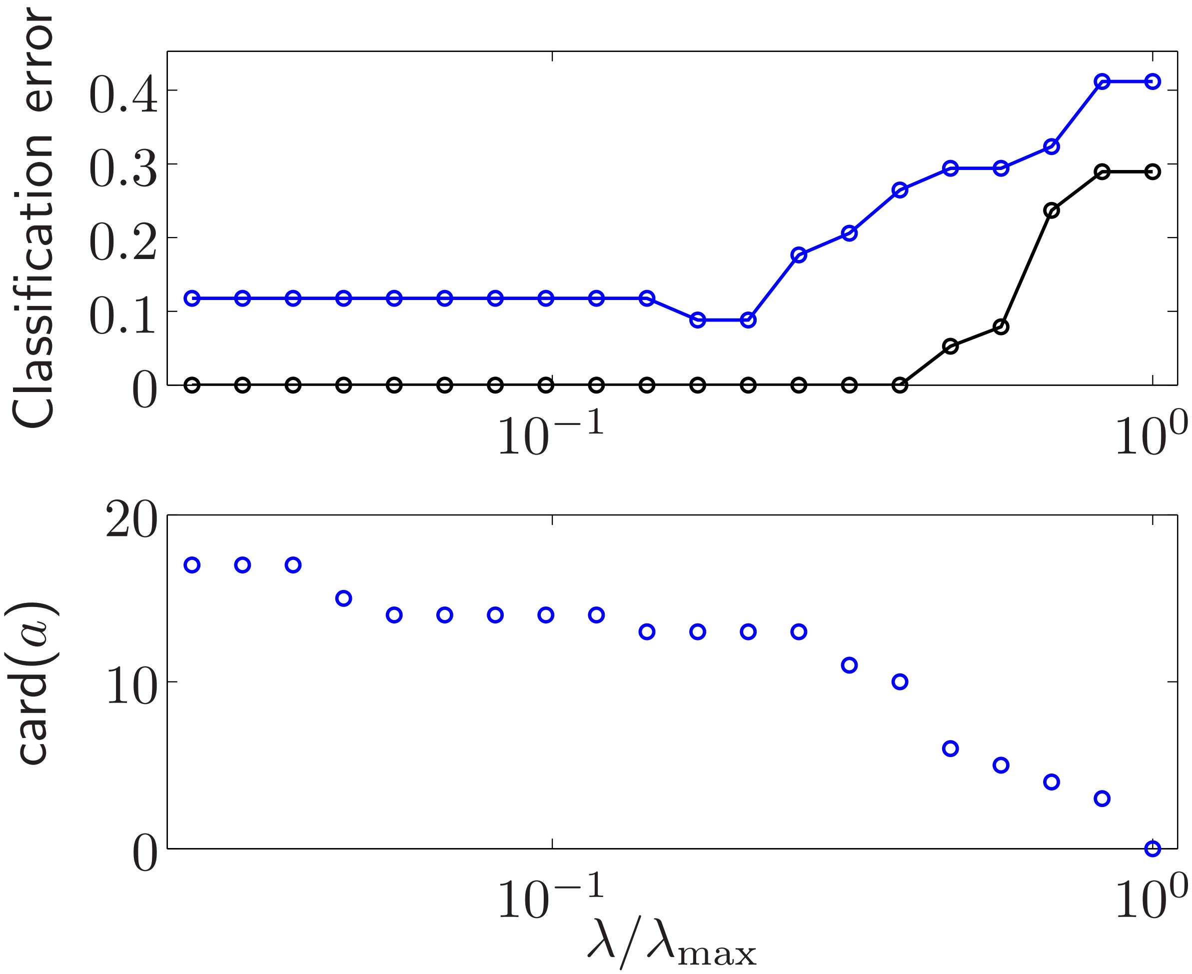


Classification error
0.4
0.3
0.2
0.1
0
10^{-3}
10^{-2}
10^{-1}
10^{0}
card(a)
60
40
20
0
10^{-3}
10^{-2}
10^{-1}
10^{0}
$\lambda/\lambda_{\max}$

Leukemia example

- taken from Golub et al, *Science* 1999
- $n = 7129$ features (gene expression data)
- $m = 72$ examples (acute leukemia patients), divided into training set (38) and validation set (34)
- outcome: type of cancer (ALL or AML)
- ℓ_1-regularized logistic regression model found using training set; classification performance checked on validation set

Results



Classification error
0.4
0.3
0.2
0.1
0
10^{-1}
10^{0}
card(a)
20
10
0
10^{-1}
10^{0}
$\lambda/\lambda_{\max}$