A Minimax Theorem with Applications to Machine Learning, Signal Processing, and Finance

Seung Jean Kim **Stephen Boyd** Alessandro Magnani

MIT ORC Seminar 12/8/05

Outline

- A minimax theorem
- Robust Fisher discriminant analysis
- Robust matched filtering
- An extension
- Worst-case Sharpe ratio maximization

A fractional function

$$f(a, B, x) = \frac{a^T x}{\sqrt{x^T B x}}, \qquad a, x \in \mathbf{R}^n, \quad B \in \mathbf{S}^n_{++}$$

- f is homogeneous, quasiconcave in x provided $a^T x \geq 0$: for $\gamma \geq 0$,

$$\{x \mid f(a, B, x) \geq \gamma\} = \left\{x \mid \gamma\sqrt{x^T B x} \leq a^T x\right\}$$

  is convex (a second-order cone)

- f is quasiconvex in (a, B) provided $a^T x \geq 0$: for $\gamma \geq 0$,

$$\{(a, B) \mid f(a, B, x) \leq \gamma\} = \left\{(a, B) \mid \gamma\sqrt{x^T B x} \geq a^T x\right\}$$

  is convex (since $\sqrt{x^T B x}$ is concave in B)

- f maximized over x by $x^\star = B^{-1} a$; optimal value $\sqrt{a^T B^{-1} a}$

A Rayleigh quotient

$f(a, B, x)^2$ is Rayleigh quotient of (aa^T, B) evaluated at x:

$$f(a, B, x)^2 = \frac{(x^T a)^2}{x^T B x}$$

- convex in (a, B)
- not quasiconcave in x

Interpretation via Gaussian

- suppose $z \sim \mathcal{N}(a, B)$, and $x \in \mathbf{R}^n$

- $z^T x \sim \mathcal{N}(a^T x, x^T B x)$, so $\mathbf{Prob}(z^T x \geq 0) = \Phi\left(\frac{a^T x}{\sqrt{x^T B x}}\right)$

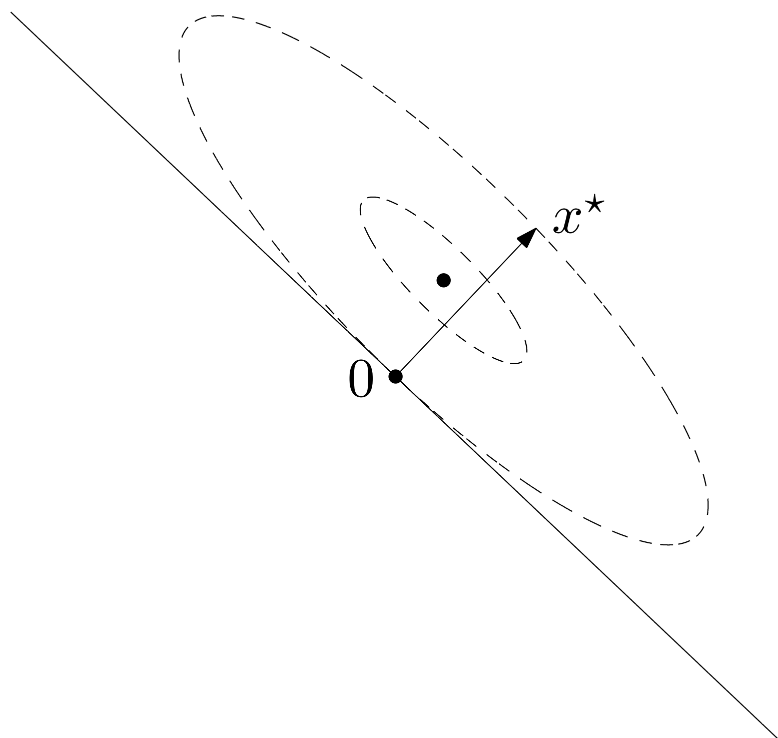
- $x^\star = B^{-1} a$ gives hyperplane through origin that maximizes probability of z being on one side

- maximum probability is $\Phi\left(\sqrt{a^T B^{-1} a}\right)$

- $a^T B^{-1} a$ measures how well a linear function can discriminate z from zero

A picture

- expand confidence ellipsoid until it touches origin
- tangent is plane that maximizes $\mathbf{Prob}(z^T x \geq 0)$



Interpretation via Chebyshev bound

- suppose $\mathbf{E}\,z = a$, $\mathbf{E}(z-a)^T(z-a) = B$ (otherwise arbitrary)
- $\mathbf{E}\,z^Tx = a^Tx$, $\mathbf{E}(z^Tx - a^Tx)^2 = x^TBx$, so by Chebyshev bound

$$\mathbf{Prob}(z^Tx \geq 0) \geq \Psi\left(\frac{a^Tx}{\sqrt{x^TBx}}\right), \qquad \Psi(u) = \frac{u_+^2}{1+u_+^2}$$

  ψ is increasing (bound is sharp)

- $x^\star = B^{-1}a$ gives hyperplane through origin that maximizes Chebyshev lower bound for $\mathbf{Prob}(z^Tx \geq 0)$
- maximum value of Chebyshev lower bound is $\dfrac{a^TB^{-1}a}{1+a^TB^{-1}a}$

Worst-case discrimination probability analysis

- **uncertain statistics**: $(a, B) \in \mathcal{U} \subseteq \mathbf{R}^n \backslash \{0\} \times \mathbf{S}^n_{++}$ (convex and compact)
- for fixed x, find **worst-case statistics**:

$$\begin{array}{ll} \text{minimize} & \mathbf{Prob}(z^T x > 0) \\ \text{subject to} & (a, B) \in \mathcal{U} \end{array}$$

where $z \sim \mathcal{N}(a, B)$

Worst-case discrimination probability analysis

- worst-case discrimination probability analysis is equivalent to

$$\begin{array}{ll} \text{minimize} & \dfrac{a^T x}{\sqrt{x^T B x}} \\ \text{subject to} & (a, B) \in \mathcal{U} \end{array}$$

- if optimal value is positive, *i.e.*, $a^T x > 0$ for all $(a, B) \in \mathcal{U}$, equivalent to convex problem

$$\begin{array}{ll} \text{minimize} & \dfrac{(a^T x)^2}{x^T B x} \\ \text{subject to} & (a, B) \in \mathcal{U} \end{array}$$

- if optimal value is negative, can solve via bisection, convex optimization

Worst-case discrimination probability maximization

- find x that maximizes worst-case discrimination probability

$$\min_{(a,B)\in\mathcal{U}} \mathbf{Prob}(z^T x \geq 0)$$

- equivalent to finding x that maximizes

$$\min_{(a,B)\in\mathcal{U}} \frac{a^T x}{\sqrt{x^T B x}}$$

- not concave in x; not clear how to solve

- studied in context of robust signal processing in 1980s
(Poor, Verdú, . . .) for some specific $\mathcal{U}$s

Weak minimax property

$$\max_{x \neq 0} \min_{(a,B)\in\mathcal{U}} \frac{x^T a}{\sqrt{x^T B x}} \leq \min_{(a,B)\in\mathcal{U}} \max_{x \neq 0} \frac{x^T a}{\sqrt{x^T B x}}$$

- LHS: worst-case discrimination probability maximization problem
- RHS can be evaluated via **convex optimization**:

$$\min_{(a,B)\in\mathcal{U}} \max_{x \neq 0} \frac{x^T a}{\sqrt{x^T B x}} = \min_{(a,B)\in\mathcal{U}} \sqrt{a^T B^{-1} a} = \left[\min_{(a,B)\in\mathcal{U}} a^T B^{-1} a \right]^{1/2}$$

- finds statistics hardest to discriminate from zero

Strong minimax property

in fact, we have **equality**:

$$\max_{x \neq 0} \min_{(a,B) \in \mathcal{U}} \frac{x^T a}{\sqrt{x^T B x}} = \min_{(a,B) \in \mathcal{U}} \max_{x \neq 0} \frac{x^T a}{\sqrt{x^T B x}}$$

(and common value is ≥ 0)

equivalently, with $z \sim \mathcal{N}(a, B)$,

$$\max_{x \neq 0} \min_{(a,B) \in \mathcal{U}} \mathbf{Prob}(z^T x \geq 0) = \min_{(a,B) \in \mathcal{U}} \max_{x \neq 0} \mathbf{Prob}(z^T x \geq 0)$$

(and common value is $\geq 1/2$)

Computing saddle point via convex optimization

- find **least favorable** statistics $(a^\star, B^\star)$, *i.e.*, minimize $a^T B^{-1} a$ over $(a, B) \in \mathcal{U}$

- $(x^\star, a^\star, B^\star)$ with $x^\star = B^{\star -1} a^\star$ is saddle point:

$$\frac{x^T a^\star}{\sqrt{x^T B^\star x}} \le \frac{x^{\star T} a^\star}{\sqrt{x^{\star T} B^\star x^\star}} \le \frac{x^{\star T} a}{\sqrt{x^{\star T} B x^\star}}, \quad \forall\, x \ne 0, \quad \forall\, (a, B) \in \mathcal{U}$$

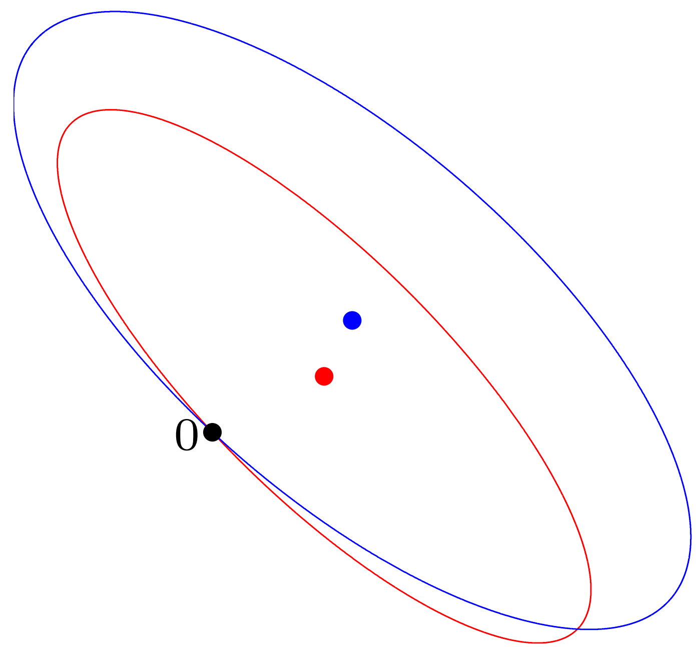
- $x^\star$ solves worst-case discrimination probability maximization problem

- $(a^\star, B^\star)$ is worst-case statistics for $x^\star$

Interpretation of least favorable statistics

find statistics with minimum Mahalanobis distance between mean and zero:

$$a^{\star T} B^{\star -1} a^{\star} = \min_{(a,B)\in\mathcal{U}} a^T B^{-1} a$$



Proof via convex analysis

- let $(a^\star, B^\star)$ be least favorable: $\min_{(a,B)\in\mathcal{U}} \sqrt{a^T B^{-1} a} = \sqrt{a^{\star T} {B^\star}^{-1} a^\star}$

- by minimax inequality

$$\max_{x\neq 0} \min_{(a,B)\in\mathcal{U}} \frac{x^T a}{\sqrt{x^T B x}} \leq \min_{(a,B)\in\mathcal{U}} \max_{x\neq 0} \frac{x^T a}{\sqrt{x^T B x}} = \sqrt{a^{\star T} {B^\star}^{-1} a^\star}$$

- $x^\star = {B^\star}^{-1} a^\star$ gives lower bound on $\sqrt{a^{\star T} {B^\star}^{-1} a^\star}$:

$$\min_{(a,B)\in\mathcal{U}} \frac{x^{\star T} a}{\sqrt{x^{\star T} B x^\star}} \leq \max_{x\neq 0} \min_{(a,B)\in\mathcal{U}} \frac{x^T a}{\sqrt{x^T B x}} \leq \sqrt{a^{\star T} {B^\star}^{-1} a^\star}$$

- if $\min_{(a,B)\in\mathcal{U}} \dfrac{x^{\star T} a}{\sqrt{x^{\star T} B x^\star}} = \dfrac{x^{\star T} a^\star}{\sqrt{x^{\star T} B^\star x^\star}} = \sqrt{a^{\star T} {B^\star}^{-1} a^\star}$, then minimax equality must hold

- $(a^\star, B^\star)$ is optimal for $\min_{(a,B)\in\mathcal{U}} a^T B^{-1} a$, so

$$2x^{\star T}(a - a^\star) - x^{\star T}(B - B^\star)x^\star \geq 0, \quad \forall\, (a, B) \in \mathcal{U}$$

  with $x^\star = B^{\star -1} a^\star$

  (since $z^\star$ minimizes convex differentiable fct. f over convex set C iff $\nabla f(z^\star)^T(z - z^\star) \geq 0\ \forall z \in C$)

- we conclude $(a^\star, B^\star)$ is optimal for $\min_{(a,B)\in\mathcal{U}} (x^{\star T} a)^2 / x^{\star T} B x^\star$, since its optimality condition is also

$$2x^{\star T}(a - a^\star) - x^{\star T}(B - B^\star)x^\star \geq 0, \quad \forall\, (a, B) \in \mathcal{U}$$

- $(a^\star, B^\star)$ must be optimal for $\min_{(a,B)\in\mathcal{U}} x^{\star T} a / (x^{\star T} B x^\star)^{1/2}$, since $\min_{(a,B)\in\mathcal{U}} x^{\star T} a / (x^{\star T} B x^\star)^{1/2} > 0$

Fisher linear discriminant analysis

- $x \sim \mathcal{N}(\mu_x, \Sigma_x)$, $y \sim \mathcal{N}(\mu_y, \Sigma_y)$
- find w that maximizes $\mathbf{Prob}(w^T x > w^T y)$
- $x - y \sim \mathcal{N}(\mu_x - \mu_y, \Sigma_x + \Sigma_y)$, so

$$\mathbf{Prob}(w^T x > w^T y) = \Phi\left(\frac{w^T(\mu_x - \mu_y)}{\sqrt{w^T(\Sigma_x + \Sigma_y)w}}\right)$$

- equivalent to maximizing $\dfrac{w^T(\mu_x - \mu_y)}{\sqrt{w^T(\Sigma_x + \Sigma_y)w}} = f(\mu_x - \mu_y, \Sigma_x + \Sigma_y, w)$
- proposed by Fisher in 1930s

Worst-case Fisher linear discrimination

- uncertain statistics:

$$(\mu_x, \mu_y, \Sigma_x, \Sigma_y) \in \mathcal{V} \subseteq \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{S}^n_{++} \times \mathbf{S}^n_{++}$$

  convex and compact; $\mu_x \neq \mu_y$ for each $(\mu_x, \mu_y, \Sigma_x, \Sigma_y) \in \mathcal{V}$

- worst-case discrimination probability (for fixed w):

$$\min_{(\mu_x, \mu_y, \Sigma_x, \Sigma_y) \in \mathcal{V}} \mathbf{Prob}(w^T x > w^T y)$$

- can compute via convex optimization

Robust Fisher linear discriminant analysis

- *worst-case discrimination probability maximization*:
  find w that maximizes worst-case discrimination probability

$$\min_{(\mu_x,\mu_y,\Sigma_x,\Sigma_y)\in\mathcal{V}} \mathbf{Prob}(w^T x > w^T y)$$

- equivalent to maximizing (over w)

$$\min_{(\mu_x,\mu_y,\Sigma_x,\Sigma_y)\in\mathcal{V}} \frac{w^T(\mu_x - \mu_y)}{\sqrt{w^T(\Sigma_x + \Sigma_y)w}}$$

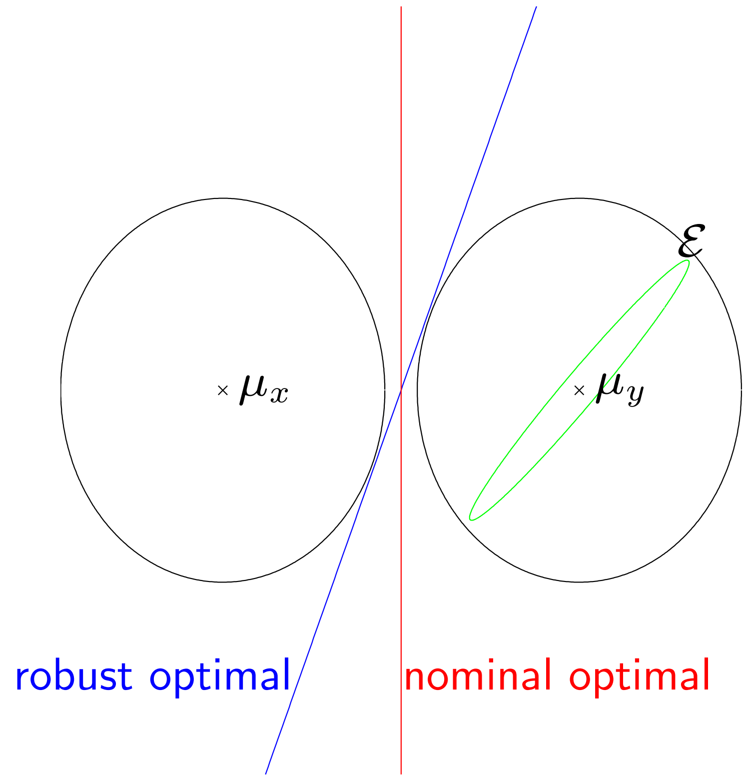
- our max-min problem, with

$$a = \mu_x - \mu_y, \qquad B = \Sigma_x + \Sigma_y$$

$$\mathcal{U} = \{(\mu_x - \mu_y, \Sigma_x + \Sigma_y) \mid (\mu_x, \mu_y, \Sigma_x, \Sigma_y) \in \mathcal{V}\}$$

Example

$x, y \in \mathbf{R}^2$; only μ_y is uncertain, $\mu_y \in \mathcal{E}$



Results

	$\mathbf{P}_{\rm nom}$	$\mathbf{P}_{\rm wc}$
nominal optimal	0.99	0.87
robust optimal	0.98	0.91

- $\mathbf{P}_{\rm nom}$: nominal discrimination probability
- $\mathbf{P}_{\rm wc}$: worst-case discrimination probability

Matched filtering

$$y(t) = s(t)a + v(t) \in \mathbf{R}^n$$

- $s(t) \in \mathbf{R}$ is desired signal; $y(t) \in \mathbf{R}^n$ is received signal
- $v(t) \sim \mathcal{N}(0, \Sigma)$ is noise
- filtered output with weight vector $w \in \mathbf{R}^n$:

$$z(t) = w^T y(t) = s(t) w^T a + w^T v(t)$$

- standard matched filtering:

  choose w to maximize (amplitude) signal to noise ratio (SNR) $\dfrac{w^T a}{\sqrt{w^T \Sigma w}}$

Robust matched filtering

- uncertain data: $(a, \Sigma) \in \mathcal{U} \subseteq \mathbf{R}^n \backslash \{0\} \times \mathbf{S}^n_{++}$ (convex and compact)
- worst-case SNR for fixed w: $\min_{(a,\Sigma)\in\mathcal{U}} \dfrac{w^T a}{\sqrt{w^T \Sigma w}}$
- **robust matched filtering**: find w that maximizes worst-case SNR
- can be solved via convex optimization & minimax theorem for $x^T a / \sqrt{x^T B x}$

Example

- $a = (2,3,2,2)$ is fixed (no uncertainty)

- uncertain Σ has form

$$\begin{bmatrix} 1 & - & + & - \\ & 1 & ? & + \\ & & 1 & ? \\ & & & 1 \end{bmatrix}, \qquad \begin{array}{lll} + & \text{means} & \Sigma_{ij} \in [0,1] \\ - & \text{means} & \Sigma_{ij} \in [-1,0] \\ ? & \text{means} & \Sigma_{ij} \in [-1,1] \end{array}$$

(and of course $\Sigma \succ 0$)

- we take 'nominal' noise covariance as $\bar{\Sigma} = \begin{bmatrix} 1 & -.5 & .5 & -.5 \\ & 1 & 0 & .5 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}$

Results

	nominal SNR	worst-case SNR
nominal optimal	5.5	3.0
robust optimal	4.9	3.6

least favorable covariance: $$\begin{bmatrix} 1 & 0 & .38 & -.12 \\ & 1 & .41 & .74 \\ & & 1 & .23 \\ & & & 1 \end{bmatrix}$$

Extension of minimax theorem

$\mathcal{U}$ convex, compact; $\mathcal{X}$ convex cone

$$\max_{x \in \mathcal{X}} \min_{(a,B) \in \mathcal{U}} \frac{a^T x}{\sqrt{x^T B x}} = \min_{(a,B) \in \mathcal{U}} \max_{w \in \mathcal{X}} \frac{a^T x}{\sqrt{x^T B x}}$$

provided there exists $\bar{x} \in \mathcal{X}$ s.t. $a^T \bar{x} > 0$ for all a with $(a, B) \in \mathcal{U}$, $i.e.$, LHS > 0

Computing saddle point via convex optimization

- convexity of min-max problem:

$$\begin{aligned}\min_{(a,B)\in\mathcal{U}} \max_{x\in\mathcal{X}} \frac{a^T x}{\sqrt{x^T B x}} &= \min_{(a,B)\in\mathcal{U}} \min_{\lambda\in\mathcal{X}^*} \left[(a+\lambda)^T B^{-1}(a+\lambda)\right]^{1/2} \\ &= \left[\min_{(a,B)\in\mathcal{U},\ \lambda\in\mathcal{X}^*} (a+\lambda)^T B^{-1}(a+\lambda)\right]^{1/2}\end{aligned}$$

  $\mathcal{X}^*$ is dual cone

- if $(a^\star, B^\star, \lambda^\star)$ is optimal for min-max problem, $(x^\star, a^\star, B^\star)$ with $x^\star = B^{\star -1}(a^\star + \lambda^\star)$ is a saddle point:

$$\frac{x^T a^\star}{\sqrt{x^T B^\star x}} \le \frac{x^{\star T} a^\star}{\sqrt{x^{\star T} B^\star x^\star}} \le \frac{x^{\star T} a}{\sqrt{x^{\star T} B x^\star}}, \quad \forall\, x \in \mathcal{X}, \quad \forall\, (a, B) \in \mathcal{U}$$

A key lemma

$$\max_{x \in \mathcal{X}} \frac{(a^T x)^2}{x^T B x} = \min_{\lambda \in \mathcal{X}^*} (a+\lambda)^T B^{-1} (a+\lambda)$$

using homogeneity of $(a^T x)^2 / x^T B x$, write LHS as

$$\max_{x \in \mathcal{X}} \frac{(a^T x)^2}{x^T B x} = \left[\min_{x \in \mathcal{X},\ a^T x = 1} x^T B x \right]^{-1}$$

by convex duality

$$\min_{x \in \mathcal{X},\ a^T x = 1} x^T B x = \max_{\bar{\lambda} \in \mathcal{X}^*,\ \mu \in \mathbf{R}} \left[-(1/4)(\bar{\lambda} + \mu a)^T B^{-1} (\bar{\lambda} + \mu a) + \mu \right]$$

defining $\lambda = \bar{\lambda}/\mu$ and optimizing over μ, RHS becomes

$$\min_{x \in \mathcal{X},\ a^T x = 1} x^T B x = \max_{\lambda \in \mathcal{X}^*} \frac{1}{(a+\lambda)^T B^{-1} (a+\lambda)}$$

Asset allocation

- n risky assets with (single period) returns $a \sim \mathcal{N}(\mu, \Sigma)$
- portfolio $w \in \mathcal{W}$ (convex); $\mathbf{1}^T w = 1$ for all $w \in \mathcal{W}$
- $w^T a \sim \mathcal{N}(w^T \mu, w^T \Sigma w)$, so probability of beating risk-free asset with return μ_{rf} is
$$\mathbf{Prob}(a^T w > \mu_{\mathrm{rf}}) = \Phi\left(\frac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}}\right)$$
- maximized by $w \in \mathcal{W}$ that maximizes *Sharpe ratio* $\dfrac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}}$ (called *tangency portfolio* w_{tp})
- we're only interested in case when maximum Sharpe ratio ≥ 0

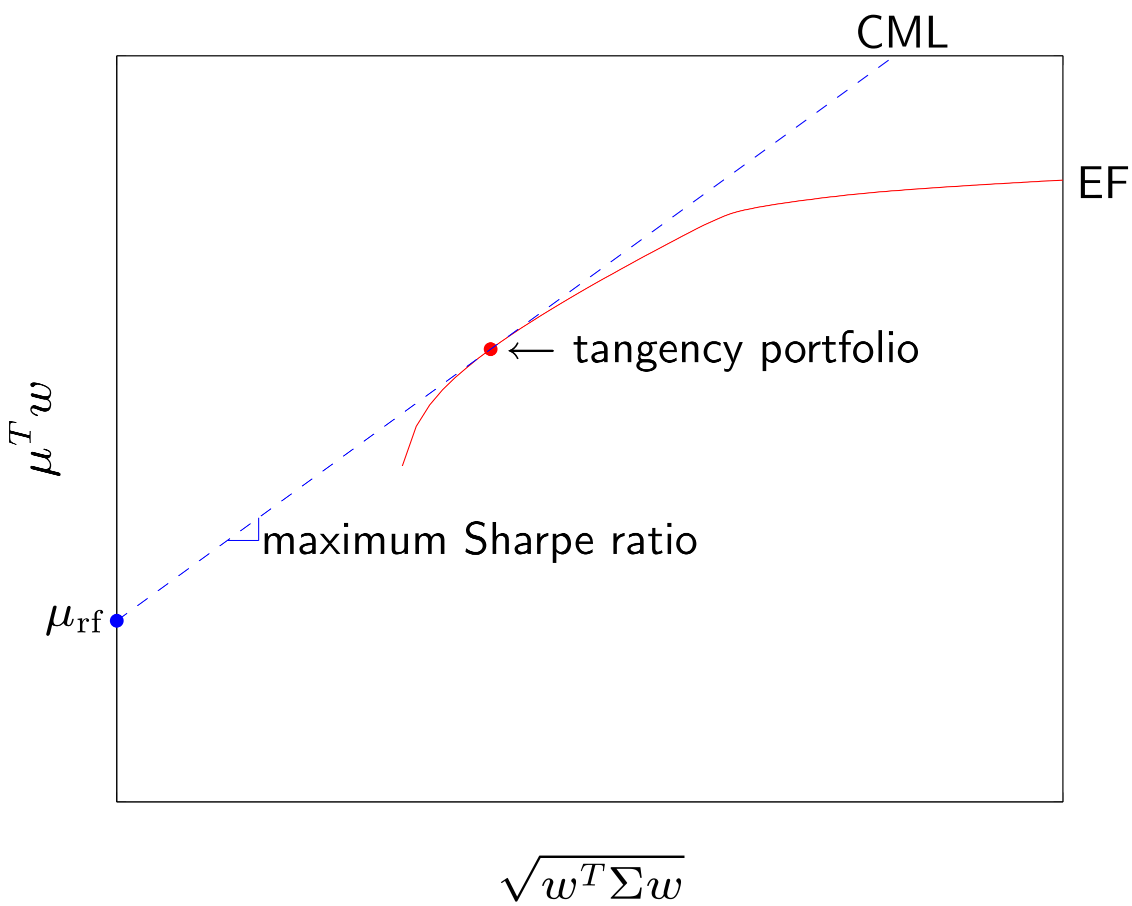
Interpretation via Chebyshev bound

- suppose $\mathbf{E}\ a = \mu$, $\mathbf{E}\,(a-\mu)^T(a-\mu) = \Sigma$ (otherwise arbitrary)
- $\mathbf{E}\ a^T w = \mu^T x$, $\mathbf{E}\,(a^T x - \mu^T x)^2 = w^T \Sigma w$, so by Chebyshev bound,

$$\mathbf{Prob}(a^T w \geq \mu_{\rm rf}) \geq \Psi\left(\frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T \Sigma w}}\right), \qquad \Psi(u) = \frac{u_+^2}{1+u_+^2}$$

  increasing (bound is tight)

- maximized by tangency portfolio



EF: efficient frontier; CML: capital market line

Asset allocation with a risk-free asset

- affine combination of risk-free asset and risky portfolio $w \in \mathcal{W}$:

$$x = \begin{bmatrix} (1-\theta)w \\ \theta \end{bmatrix}$$

  θ is amount of risk-free asset

- Tobin's two-fund separation theorem:
  - risk s and return r of any $x = ((1-\theta)w, \theta)$ with $w \in \mathcal{W}$ cannot lie above **capital market line (CML)**

$$r = \mu_{\rm rf} + S^\star s, \qquad S^\star = \max_{w \in \mathcal{W}} \frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T \Sigma w}}$$

  - CML is formed by affine combinations of two portfolios: tangency portfolio $(w_{\rm tp}, 0)$ and risk-free portfolio $(0, 1)$

Worst-case Sharpe ratio

- uncertain statistics: $(\mu, \Sigma) \in \mathcal{V} \subseteq \mathbf{R}^n \times \mathbf{S}^n_{++}$ (convex and compact)

- for fixed $w \in \mathcal{W}$, find worst-case statistics by computing worst-case SR

$$\min_{(\mu,\Sigma)\in\mathcal{V}} \frac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}} = \min_{(\mu,\Sigma)\in\mathcal{V}} \frac{(\mu - \mu_{\mathrm{rf}}\mathbf{1})^T w}{\sqrt{w^T \Sigma w}}$$

. . . can be computed via convex optimization

- related to worst-case probability of beating risk-free asset

$$\min_{(\mu,\Sigma)\in\mathcal{V}} \mathbf{Prob}(a^T w > \mu_{\mathrm{rf}}) = \Phi\left(\min_{(\mu,\Sigma)\in\mathcal{V}} \frac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}}\right)$$

Worst-case Sharpe ratio maximization

find portfolio that maximizes worst-case SR:

$$\begin{array}{ll} \text{maximize} & \displaystyle\min_{(\mu,\Sigma)\in\mathcal{V}} \frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T\Sigma w}} \\ \text{subject to} & w \in \mathcal{W} \end{array}$$

this portfolio (called *robust tangency portfolio* $w_{\rm rtp}$) maximizes

- worst-case probability of beating risk-free asset

$$\min_{(\mu,\Sigma)\in\mathcal{V}} \mathbf{Prob}(w^T a > \mu_{\rm rf}), \qquad a \sim \mathcal{N}(\mu, \Sigma)$$

- worst-case Chebyshev lower bound for $\mathbf{Prob}(a^T w \geq \mu_{\rm rf})$

$$\min_{(\mu,\Sigma)\in\mathcal{V}} \inf\{\mathbf{Prob}(a^T w \geq \mu_{\rm rf}) \mid \mathbf{E}\, a = \mu,\ \mathbf{E}\,(a-\mu)^T(a-\mu) = \Sigma\}$$

Solution via minimax

$\mathbf{1}^T w = 1$ for all $w \in \mathcal{W}$ so

$$\frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T \Sigma w}} = \frac{(\mu - \mu_{\rm rf}\mathbf{1})^T w}{\sqrt{w^T \Sigma w}}, \qquad \forall\, w \in \mathcal{W}$$

can use minimax theorem for $a^T x / \sqrt{x^T B x}$, with

$$a = \mu - \mu_{\rm rf}\mathbf{1}, \qquad B = \Sigma, \qquad \mathcal{X} = \mathbf{R}\mathcal{W}$$

$$\mathcal{U} = \{(\mu - \mu_{\rm rf}\mathbf{1}, \Sigma) \mid (\mu, \Sigma) \in \mathcal{V}\}$$

Minimax result for Sharpe ratio

suppose there exists portfolio $\bar{w}$ s.t. $\mu^T w > \mu_{\rm rf}$ for all $(\mu, \Sigma) \in \mathcal{V}$

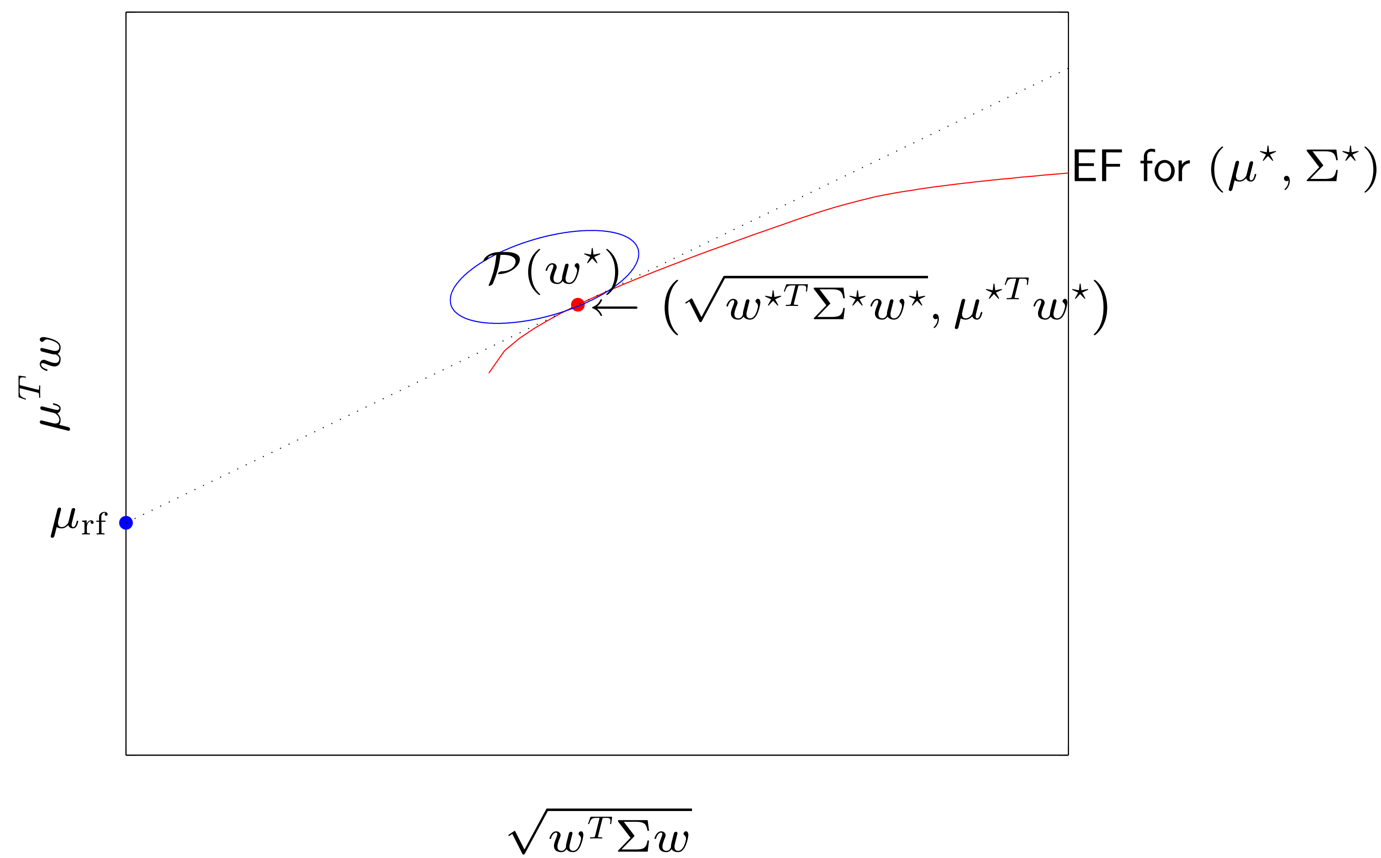
- strong minimax property:

$$0 < \max_{w \in \mathcal{W}} \min_{(\mu,\Sigma)\in\mathcal{V}} \frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T \Sigma w}} = \min_{(\mu,\Sigma)\in\mathcal{V}} \max_{w \in \mathcal{W}} \frac{\mu^T w - \mu_{\rm rf}}{\sqrt{w^T \Sigma w}}$$

- saddle point $(w^\star, \mu^\star, \Sigma^\star)$ can be computed via **convex optimization**

$$\frac{\mu^{\star T} w - \mu_{\rm rf}}{\sqrt{w^T \Sigma^\star w}} \le \frac{\mu^{\star T} w^\star - \mu_{\rm rf}}{\sqrt{w^{\star T} \Sigma^\star w^\star}} \le \frac{\mu^T w^\star - \mu_{\rm rf}}{\sqrt{w^{\star T} \Sigma w^\star}}, \quad \forall\, w \in \mathcal{W}, \quad \forall\, (\mu, \Sigma) \in \mathcal{V}$$

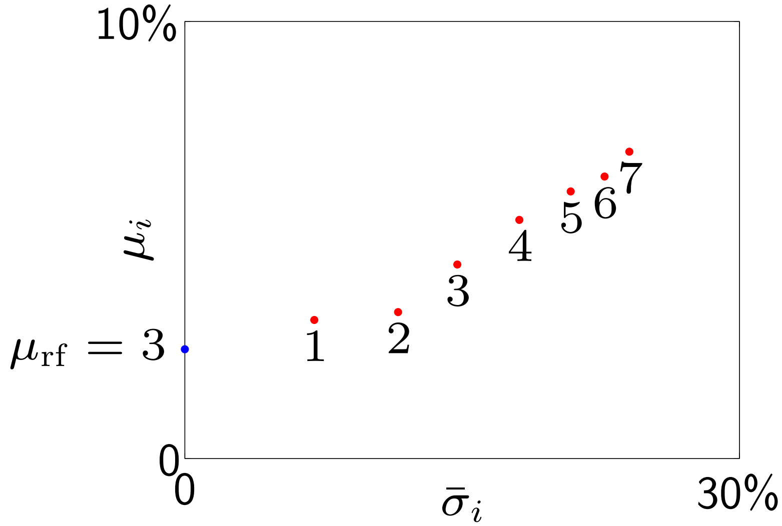
Saddle-point property



$\mathcal{P}(w^\star) = \left\{ (\mu^T w^\star, \sqrt{w^{\star T}\Sigma w^\star}) \mid (\mu, \Sigma) \in \mathcal{V} \right\}$ is **risk-return set** of $w^\star$

Example

7 risky assets with nominal returns $\bar{\mu}_i$, variances $\bar{\sigma}_i^2$, correlation matrix $\bar{\Omega}$



$$\bar{\Omega} = \begin{bmatrix} 1.00 & .07 & -.12 & .43 & -.11 & .24 & .25 \\ & 1.00 & .73 & -.14 & .09 & .28 & -.10 \\ & & 1.00 & .14 & .5 & .52 & -.13 \\ & & & 1.00 & .04 & .35 & .38 \\ & & & & 1.00 & .7 & .04 \\ & & & & & 1.00 & -.09 \\ & & & & & & 1.00 \end{bmatrix}$$

Example

- total short position is limited to 30% of total long position
- mean uncertainty model:

$$|\mathbf{1}^T\mu - \mathbf{1}^T\bar{\mu}| \leq 0.1|\mathbf{1}^T\bar{\mu}|, \qquad |\mu_i - \bar{\mu}_i| \leq 0.2|\bar{\mu}_i|, \quad i = 1, \ldots, 6$$

- covariance uncertainty model:

$$\|\Sigma - \bar{\Sigma}\|_F \leq 0.1\|\bar{\Sigma}\|_F, \qquad |\Sigma_{ij} - \bar{\Sigma}_{ij}| \leq 0.2|\bar{\Sigma}_{ij}|, \quad i, j = 1, \ldots, 6$$

$\bar{\Sigma}$ is nominal covariance

Results

	nominal SR	worst-case SR
nominal MP	1.56	0.56
robust MP	1.23	0.77

	$\mathbf{P}_{\rm nom}$	$\mathbf{P}_{\rm wc}$
nominal MP	0.92	0.77
robust MP	0.89	0.71

- $\mathbf{P}_{\rm nom}$: probability of beating risk-free asset with nominal statistics $(\bar{\mu}, \bar{\Sigma})$
- $\mathbf{P}_{\rm wc}$: probability of beating risk-free asset with worst-case statistics

Two-fund separation under model uncertainty

- *robust two-fund separation theorem*:
  - **risk-return set** of any $x = ((1-\theta)w, \theta)$ with $w \in \mathcal{W}$ cannot lie above **robust capital market line (RCML)**

$$r = \mu_{\mathrm{rf}} + S^\star s$$

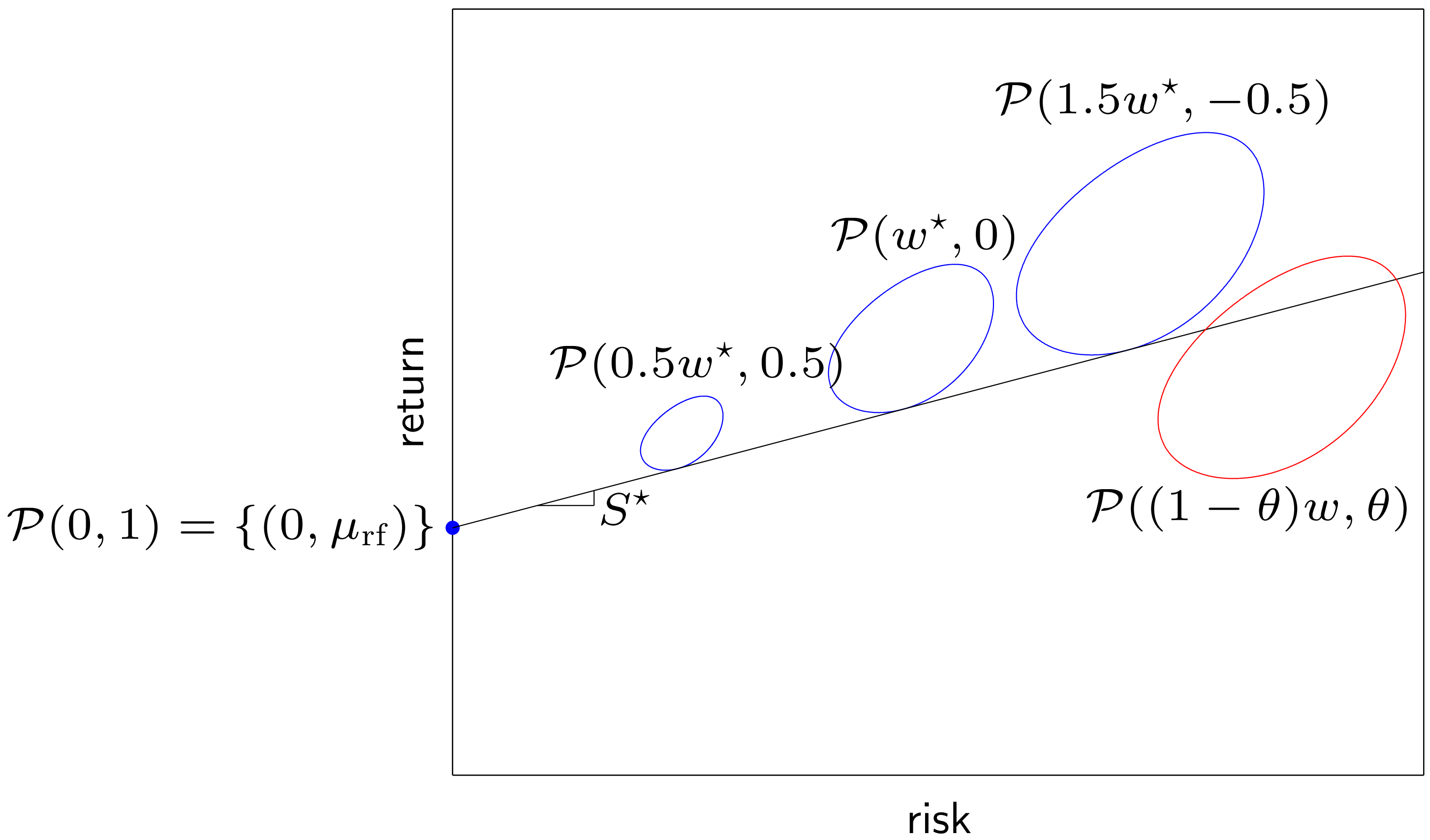
with slope $S^\star = \max_{w \in \mathcal{W}} \min_{(\mu,\Sigma)\in\mathcal{V}} \dfrac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}} = \min_{(\mu,\Sigma)\in\mathcal{V}} \max_{w \in \mathcal{W}} \dfrac{\mu^T w - \mu_{\mathrm{rf}}}{\sqrt{w^T \Sigma w}}$

  - RCML is formed by affine combinations of two portfolios:

robust tangency portfolio $(w_{\mathrm{rtp}}, 0)$, risk-free portfolio $(0, 1)$

- for $\mathcal{V} = \{(\mu, \Sigma)\}$, reduces to Tobin's two-fund separation theorem

Two-fund separation under model uncertainty



$\mathcal{P}((1-\theta)w, \theta)$ is risk-return set of $((1-\theta)w, \theta)$

Summary & comments

- the function $f(a,B,x) = \dfrac{a^T x}{\sqrt{x^T B x}}$ comes up alot

- what was known (??): simple minimax result, no constraints on x, product form for $\mathcal{U}$

- what is new (??): general minimax theorem; convex optimization method for computing saddle point

- minimax theorem for $\dfrac{\mathbf{Tr} AX}{\mathbf{Tr} BX}$ holds (with conditions)

- minimax theorem for $\dfrac{x^T A x}{x^T B x}$ generally false

References

- Kim and Boyd, Two-fund separation under model uncertainty, in preparation

- Kim, Magnani, and Boyd, Robust Fisher discriminant analysis, `www.stanford.edu/~boyd/research.html`

- Verdú and Poor, On Minimax Robustness: A General Approach and Applications, *IEEE Transactions on Information Theory*, 30(2): 328-340, 1984

- Sion, On general minimax theorems, *Pacific Journal of Mathematics*, 8(1): 171-176, 1958