Arbitrary-norm Support Vector Machine. Properties and Applications

Emilio Carrizosa
Universidad de Sevilla, Spain
http://ecarriz.us.es

Brussels, 16th of March 2005

Research project from Spanish Science Ministry

- Rafael Blanquero
- Emilio Carrizosa
- Eduardo Conde
- José Gordillo
- Belén Martín-Barragán
- Frank Plastria
- Dolores Romero Morales

Support Vector Machines: A Data Mining tool

In between many fields

- Machine Learning / Statistics
- O.R. ...

With many applications ...

- Bioinformatics (gene expression, ...) (e.g. 327 papers have the term "Support Vector Machines" in title/abstract in Bioinformatics
- CRM
- Banking (credit scoring, credit card fraud detection, ...)
- Internet (spam detection, ...)
- Pharmacology (drug design, ...)

Support Vector Machines: A Data Mining tool

In between many fields

- Machine Learning / Statistics
- O.R. . . .

With many applications . . .

- Bioinformatics (gene expression, . . .) (e.g. 327 papers have the term "Support Vector Machines" in title/abstract in Bioinformatics
- CRM
- Banking (credit scoring, credit card fraud detection, . . .)
- Internet (spam detection, . . .)
- Pharmacology (drug design, . . .)

Support Vector Machines: A Data Mining tool

In between many fields

- Machine Learning / Statistics
- O.R. . . .

With many applications . . .

- Bioinformatics (gene expression, . . .) (e.g. 327 papers have the term "Support Vector Machines" in title/abstract in Bioinformatics
- CRM
- Banking (credit scoring, credit card fraud detection, . . .)
- Internet (spam detection, . . .)
- Pharmacology (drug design, . . .)

Supervised classification (Discriminant analysis)

- Population: $\mathcal{O}$
- $u \in \mathcal{O} \hookrightarrow \begin{cases} x_u \in \mathbb{R}^N & (\text{ features }) \\ y_u \in \mathcal{C} & (\text{ class labels }) \end{cases}$
- $I \subset \mathcal{O}$: *training sample*: individuals u for which the pair (x_u, y_u) *is known*
- Aim: build a classification rule $\varphi : \mathbb{R}^N \longmapsto \mathcal{C}$, i.e., build a function assigning labels to objects u for which just x_u (and not y_u!) is known

Supervised classification (Discriminant analysis)

- Population: $\mathcal{O}$
- $u \in \mathcal{O} \hookrightarrow \begin{cases} x_u \in \mathbb{R}^N & (\text{ features }) \\ y_u \in \mathcal{C} & (\text{ class labels }) \end{cases}$
- $I \subset \mathcal{O}$: *training sample*: individuals u for which the pair (x_u, y_u) *is known*
- Aim: build a classification rule $\varphi : \mathbb{R}^N \longmapsto \mathcal{C}$, i.e., build a function assigning labels to objects u for which just x_u (and not y_u!) is known

Fisher, 1936

Classification with 2 labels

- $I \neq \emptyset$, $|I| < \infty$. $I = I_+ \cup I_-$, with $I^+ \cap I^- = \emptyset$, $I_+, I_- \neq \emptyset$.
- $i \in I \hookrightarrow \begin{cases} x_i \in \mathbb{R}^N \\ y_i := \begin{cases} 1, & \text{if } i \in I_+ \\ -1, & \text{if } i \in I_- \end{cases} \end{cases}$

For $\omega \in \mathbb{R}^N$, $\beta \in \mathbb{R}$

$$x \in \mathbb{R}^N \mapsto \begin{cases} 1, & \text{if } \omega^\top x + \beta > 0 \\ -1, & \text{if } \omega^\top x + \beta < 0 \end{cases}$$

We force $\omega \in \Omega := \left\{\omega \in \mathbb{R}^N : q_j^\top \omega \geq 0 \, \forall j \in J\right\}$, for a finite set $\{q_j : j \in J\}$

Classification with 2 labels

- $I \neq \emptyset$, $|I| < \infty$. $I = I_+ \cup I_-$, with $I^+ \cap I^- = \emptyset$, $I_+, I_- \neq \emptyset$.
- $i \in I \hookrightarrow \begin{cases} x_i \in \mathbb{R}^N \\ y_i := \begin{cases} 1, & \text{if } i \in I_+ \\ -1, & \text{if } i \in I_- \end{cases} \end{cases}$

For $\omega \in \mathbb{R}^N$, $\beta \in \mathbb{R}$

$$x \in \mathbb{R}^N \mapsto \begin{cases} 1, & \text{if } \omega^\top x + \beta > 0 \\ -1, & \text{if } \omega^\top x + \beta < 0 \end{cases}$$

We force $\omega \in \Omega := \left\{\omega \in \mathbb{R}^N : q_j^\top \omega \geq 0 \, \forall j \in J\right\}$, for a finite set $\{q_j : j \in J\}$

Classification with 2 labels

- $I \neq \emptyset$, $|I| < \infty$. $I = I_+ \cup I_-$, with $I^+ \cap I^- = \emptyset$, $I_+, I_- \neq \emptyset$.
- $i \in I \hookrightarrow \begin{cases} x_i \in \mathbb{R}^N \\ y_i := \begin{cases} 1, & \text{if } i \in I_+ \\ -1, & \text{if } i \in I_- \end{cases} \end{cases}$

For $\omega \in \mathbb{R}^N$, $\beta \in \mathbb{R}$

$$x \in \mathbb{R}^N \mapsto \begin{cases} 1, & \text{if } \omega^\top x + \beta > 0 \\ -1, & \text{if } \omega^\top x + \beta < 0 \end{cases}$$

We force $\omega \in \Omega := \left\{ \omega \in \mathbb{R}^N : \; q_j^\top \omega \geq 0 \, \forall j \in J \right\}$, for a finite set $\{q_j : j \in J\}$

Classification with 2 labels

- $I \neq \emptyset$, $|I| < \infty$. $I = I_+ \cup I_-$, with $I^+ \cap I^- = \emptyset$, $I_+, I_- \neq \emptyset$.
- $i \in I \hookrightarrow \begin{cases} x_i \in \mathbb{R}^N \\ y_i := \begin{cases} 1, & \text{if } i \in I_+ \\ -1, & \text{if } i \in I_- \end{cases} \end{cases}$

For $\omega \in \mathbb{R}^N$, $\beta \in \mathbb{R}$

$$x \in \mathbb{R}^N \mapsto \begin{cases} 1, & \text{if } \omega^\top x + \beta > 0 \\ -1, & \text{if } \omega^\top x + \beta < 0 \end{cases}$$

We force $\omega \in \Omega := \left\{\omega \in \mathbb{R}^N : \; q_j^\top \omega \geq 0 \, \forall j \in J\right\}$, for a finite set $\{q_j : j \in J\}$

Can we do it right, i.e. score 100%. . . . . . at least on I?

$(\{\omega_i : i \in I_+\}, \{\omega_i : i \in I_-\})$: *Ω-separable* if $\exists(\omega, \beta) \in \Omega \times \mathbb{R}$:
$y_i \left(\omega^\top x_i + \beta\right) > 0 \quad \forall i \in I$
Such (ω, β) : *Ω-separating* $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$.

Ω- Separability: equivalent to

$$(conv(\{x_i : i \in I_+\}) + cone(\{q_j : j \in J\})) \cap conv(\{x_i : i \in I_-\}) = \emptyset$$

If (ω, β) : Ω-separating, since, by assumption, $I_+, I_- \neq \emptyset$, one has $\omega \neq 0$.

Can we do it right, i.e. score 100%. . . . . . at least on I?

$(\{\omega_i : i \in I_+\}, \{\omega_i : i \in I_-\})$: *Ω-separable* if $\exists(\omega, \beta) \in \Omega \times \mathbb{R}$:
$y_i\left(\omega^\top x_i + \beta\right) > 0 \quad \forall i \in I$
Such (ω, β) : *Ω-separating* $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$.

Ω- Separability: equivalent to

$$(conv(\{x_i : i \in I_+\}) + cone(\{q_j : j \in J\})) \cap conv(\{x_i : i \in I_-\}) = \emptyset$$

If (ω, β) : Ω-separating, since, by assumption, $I_+, I_- \neq \emptyset$, one has $\omega \neq 0$.

Can we do it right, i.e. score 100%. . . . . . at least on I?

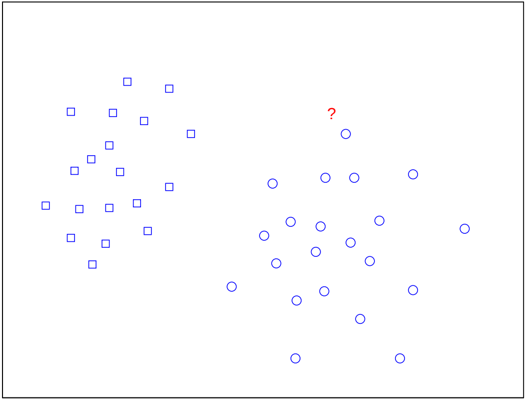
$(\{\omega_i : i \in I_+\}, \{\omega_i : i \in I_-\})$: *Ω-separable* if $\exists(\omega, \beta) \in \Omega \times \mathbb{R}$:
$y_i \left(\omega^\top x_i + \beta\right) > 0 \quad \forall i \in I$
Such (ω, β) : *Ω-separating* $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$.

Ω- Separability: equivalent to

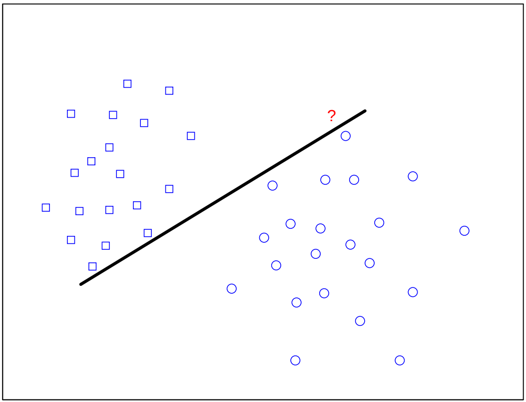
$$(conv(\{x_i : i \in I_+\}) + cone(\{q_j : j \in J\})) \cap conv(\{x_i : i \in I_-\}) = \emptyset$$

If (ω, β) : Ω-separating, since, by assumption, $I_+, I_- \neq \emptyset$, one has $\omega \neq 0$.

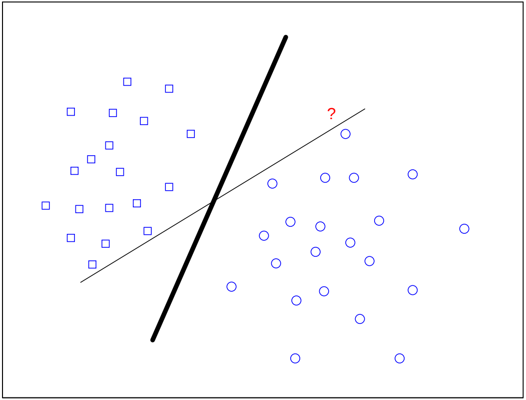
How to chose one linear classifier?



How to chose one linear classifier?



How to chose one linear classifier?



Problem statement

Halfspace of misclassification

- for $i \in I_+ : H(\omega, \beta)^- := \{x \in \mathbb{R}^N : \omega^\top x + \beta \leq 0\}$
- for $i \in I_- : H(\omega, \beta)^+ := \{x \in \mathbb{R}^N : \omega^\top x + \beta \geq 0\}$

- $\delta_i(\omega, \beta) :=$ distance (measured by some norm $\|\cdot\|$) from x_i to halfspace of misclassification
- *margin*: $\min_{i \in I} \delta_i(\omega, \beta)$

Maximize the margin

$$\max_{\omega \in \Omega \setminus \{0\}} \left(\min_{i \in I} \delta_i(\omega, \beta) \right)$$

Problem statement

Halfspace of misclassification

- for $i \in I_+ : H(\omega, \beta)^- := \{x \in \mathbb{R}^N : \omega^\top x + \beta \leq 0\}$
- for $i \in I_- : H(\omega, \beta)^+ := \{x \in \mathbb{R}^N : \omega^\top x + \beta \geq 0\}$

- $\delta_i(\omega, \beta) :=$ distance (measured by some norm $\|\cdot\|$) from x_i to halfspace of misclassification
- *margin*: $\min_{i \in I} \delta_i(\omega, \beta)$

Maximize the margin

$$\max_{\omega \in \Omega \setminus \{0\}} \left(\min_{i \in I} \delta_i(\omega, \beta) \right)$$

Problem statement

Halfspace of misclassification

- for $i \in I_+ : H(\omega, \beta)^- := \{x \in \mathbb{R}^N : \omega^\top x + \beta \leq 0\}$
- for $i \in I_- : H(\omega, \beta)^+ := \{x \in \mathbb{R}^N : \omega^\top x + \beta \geq 0\}$

- $\delta_i(\omega, \beta) :=$ distance (measured by some norm $\|\cdot\|$) from x_i to halfspace of misclassification
- *margin*: $\min_{i \in I} \delta_i(\omega, \beta)$

Maximize the margin

$$\max_{\omega \in \Omega \setminus \{0\}} \left(\min_{i \in I} \delta_i(\omega, \beta) \right)$$

$\delta_i(\omega, \beta) :=$ distance (measured by $\| \cdot \|$) from x_i to halfspace of misclassification

$$\delta_i(\omega, \beta) = \max \left\{ \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}, 0 \right\}$$

- For $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$: Ω-separable, equivalent to

$$\begin{array}{ll} \max & \min_{i \in I} \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ} \\ \text{s.t.} & \omega \in \Omega \setminus \{0\} \end{array}$$

$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \ \forall i \in I \\ & \omega \in \Omega \end{array} \qquad (P)$$

$\delta_i(\omega, \beta) :=$ distance (measured by $\|\cdot\|$) from x_i to halfspace of misclassification

$$\delta_i(\omega, \beta) = \max\left\{\frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}, 0\right\}$$

- For $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$: Ω-separable, equivalent to

$$\begin{array}{ll} \max & \min_{i \in I} \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ} \\ \text{s.t.} & \omega \in \Omega \setminus \{0\} \end{array}$$

$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \; \forall i \in I \\ & \omega \in \Omega \end{array} \qquad (P)$$

$\delta_i(\omega, \beta) :=$ distance (measured by $\|\cdot\|$) from x_i to halfspace of misclassification

$$\delta_i(\omega, \beta) = \max \left\{ \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}, 0 \right\}$$

- For $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$: Ω-separable, equivalent to

$$\begin{array}{ll} \max & \min_{i \in I} \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ} \\ \text{s.t.} & \omega \in \Omega \setminus \{0\} \end{array}$$

$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \; \forall i \in I \\ & \omega \in \Omega \end{array} \qquad (P)$$

$\delta_i(\omega, \beta) :=$ distance (measured by $\|\cdot\|$) from x_i to halfspace of misclassification

$$\delta_i(\omega, \beta) = \max\left\{\frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}, 0\right\}$$

- For $(\{x_i : i \in I_+\}, \{x_i : i \in I_-\})$: Ω-separable, equivalent to

$$\begin{array}{ll} \max & \min_{i\in I} \frac{y_i(\omega^\top x_i+\beta)}{\|\omega\|^\circ} \\ \text{s.t.} & \omega \in \Omega \setminus \{0\} \end{array}$$

$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \, \forall i \in I \\ & \omega \in \Omega \end{array} \qquad (P)$$

ℓ_2

$$\begin{array}{ll} \min & \omega^\top \omega \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \, \forall i \in I \\ & \omega \in \Omega \end{array}$$

ℓ_1

$$\begin{array}{lll} \min & t & \\ \text{s.t.} & y_i\left(\omega^\top x_i + \beta\right) \geq 1 & \forall i \in I \\ & -t \leq \omega_k \leq t & \forall k = 1, 2, \ldots, N \\ & \omega \in \Omega & \end{array}$$

ℓ_∞

$$\begin{array}{lll} \min & e^\top \omega_+ + e^\top \omega_- & \\ \text{s.t.} & y_i\left(\omega_+^\top x_i - \omega_-^\top x_i + \beta\right) \geq 1 & \forall i \in I \\ & \omega_+, \omega_- \geq 0 & \\ & \omega_+ - \omega_- \in \Omega & \end{array}$$

ℓ_2

$$\begin{array}{ll} \min & \omega^\top \omega \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \, \forall i \in I \\ & \omega \in \Omega \end{array}$$

ℓ_1

$$\begin{array}{lll} \min & t & \\ \text{s.t.} & y_i \left(\omega^\top x_i + \beta\right) \geq 1 & \forall i \in I \\ & -t \leq \omega_k \leq t & \forall k = 1, 2, \ldots, N \\ & \omega \in \Omega & \end{array}$$

ℓ_∞

$$\begin{array}{lll} \min & e^\top \omega_+ + e^\top \omega_- & \\ \text{s.t.} & y_i \left(\omega_+^\top x_i - \omega_-^\top x_i + \beta\right) \geq 1 & \forall i \in I \\ & \omega_+, \omega_- \geq 0 & \\ & \omega_+ - \omega_- \in \Omega & \end{array}$$

ℓ_2

$$
\begin{array}{ll}
\min & \omega^\top \omega \\
\text{s.t.} & y_i(\omega^\top x_i + \beta) \geq 1 \, \forall i \in I \\
& \omega \in \Omega
\end{array}
$$

ℓ_1

$$
\begin{array}{lll}
\min & t & \\
\text{s.t.} & y_i \left(\omega^\top x_i + \beta\right) \geq 1 & \forall i \in I \\
& -t \leq \omega_k \leq t & \forall k = 1, 2, \ldots, N \\
& \omega \in \Omega &
\end{array}
$$

ℓ_∞

$$
\begin{array}{lll}
\min & e^\top \omega_+ + e^\top \omega_- & \\
\text{s.t.} & y_i \left(\omega_+^\top x_i - \omega_-^\top x_i + \beta\right) \geq 1 & \forall i \in I \\
& \omega_+, \omega_- \geq 0 & \\
& \omega_+ - \omega_- \in \Omega &
\end{array}
$$

... The advantages of using polyhedral norms ...

Cost of getting info

- π_k cost associated with k-th variable
- Total cost

$$\Pi(\omega, \beta) = \sum_{k:\, \omega_k \neq 0} \pi_k$$

Objective

$$\{\max \min_i \delta_i(\omega, \beta);\ \min \Pi(\omega, \beta)\}$$

. . . The advantages of using polyhedral norms . . .

Cost of getting info

- π_k cost associated with k-th variable
- Total cost

$$\Pi(\omega, \beta) = \sum_{k:\, \omega_k \neq 0} \pi_k$$

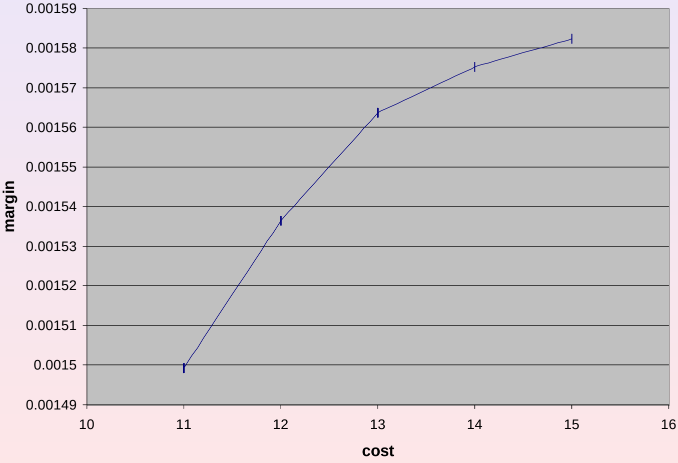
Objective

$$\{\max \min_i \delta_i(\omega, \beta);\ \min \Pi(\omega, \beta)\}$$

Formulation

$$
\begin{array}{ll}
\max & y, \quad \min \sum_{k=1}^{N} \pi_k z_k \\
\text{s.t.:} & \sum_{k=1}^{N} \phi_k(x^u)\left(\alpha^i_{+k} - \alpha^i_{-k} - \alpha^j_{+k} + \alpha^j_{-k}\right) + \beta^i - \beta^j - y \geq 0, \quad \forall i \neq ; \; u \in I_i \\
& \sum_{c=1}^{C} \sum_{k=1}^{N} \left(\alpha^c_{+k} + \alpha^c_{-k}\right) \leq 1 \\
& \sum_{h:k \preceq h} \sum_{c=1}^{C} \left(\alpha^c_{+k} + \alpha^c_{-k}\right) \leq z_k, \quad \forall k = 1, 2, \ldots, N \\
& \alpha^c_{+k} \geq 0 \quad \forall k = 1, 2, \ldots, N; \; c = 1, 2, \ldots, C \\
& \alpha^c_{-k} \geq 0 \quad \forall k = 1, 2, \ldots, N; \; c = 1, 2, \ldots, C \\
& \beta^c \geq 0 \quad \forall c = 1, 2, \ldots, C \\
& z_k \in \{0, 1\} \quad \forall k = 1, 2, \ldots, N
\end{array}
$$

Supported efficient solutions



- Separability conditions: rather strong. Still, model important since
  - Cornerstone for more sophisticated models (valid also for non-separable sets)
  - Separability may happen e.g. in small-size high-dimensional data sets
- Let's address first the Ω-separable case ...

- For $\omega \in \mathbb{R}^N \setminus \{0\}, \beta_+ \leq \beta_-$, define the band

$$\mathcal{B}(\omega, \beta_-, \beta_+) = \{x \in \mathbb{R}^N : \omega^\top x + \beta_+ \leq 0 \leq \omega^\top x + \beta_-\}$$

- Its width $W(\mathcal{B}(\omega, \beta_-, \beta_+))$:

$$\begin{aligned} W(\mathcal{B}(\omega, \beta_-, \beta_+)) &= \min \quad \|x_+ - x_-\| \\ &\quad \text{s.t.} \quad \begin{array}{lcl} x_- & \in & H(\omega, \beta_-)^- \\ x_+ & \in & H(\omega, \beta_+)^+ \end{array} \\ &= \frac{\beta_- - \beta_+}{\|\omega\|^\circ} \end{aligned}$$

- For $\omega \in \mathbb{R}^N \setminus \{0\}, \beta_+ \leq \beta_-$, define the band

$$\mathcal{B}(\omega, \beta_-, \beta_+) = \left\{x \in \mathbb{R}^N : \omega^\top x + \beta_+ \leq 0 \leq \omega^\top x + \beta_-\right\}$$

- Its width $W(\mathcal{B}(\omega, \beta_-, \beta_+))$:

$$\begin{array}{rcll} W(\mathcal{B}(\omega, \beta_-, \beta_+)) & = & \min & \|x_+ - x_-\| \\ & & \text{s.t.} & \begin{array}{lcl} x_- & \in & H(\omega, \beta_-)^- \\ x_+ & \in & H(\omega, \beta_+)^+ \end{array} \\ & = & \frac{\beta_- \beta_+}{\|\omega\|^\circ} & \end{array}$$

- For $\omega \in \mathbb{R}^N \setminus \{0\}, \beta_+ \leq \beta_-$, define the band

$$\mathcal{B}(\omega, \beta_-, \beta_+) = \left\{x \in \mathbb{R}^N : \omega^\top x + \beta_+ \leq 0 \leq \omega^\top x + \beta_-\right\}$$

- Its width $W(\mathcal{B}(\omega, \beta_-, \beta_+))$:

$$\begin{array}{rcll} W(\mathcal{B}(\omega, \beta_-, \beta_+)) & = & \min & \|x_+ - x_-\| \\ & & \text{s.t.} & \begin{array}{rcl} x_- & \in & H(\omega, \beta_-)^- \\ x_+ & \in & H(\omega, \beta_+)^+ \end{array} \\ & = & \frac{\beta_- \beta_+}{\|\omega\|^\circ} & \end{array}$$

$$W(\mathcal{B}(\omega, \beta_-, \beta_+)) = \frac{\beta_- \beta_+}{\|\omega\|^\circ}$$

$$\begin{array}{ll} \max & W(\mathcal{B}(\omega, \beta_-, \beta_+)) \\ \text{s.t.} & \{x_i : i \in I_-\} \subset H(\omega, \beta_-)^- \\ & \{x_i : i \in I_+\} \subset H(\omega, \beta_+)^+ \\ & \omega \in \Omega \end{array}$$

$$\begin{array}{lll} \max & \frac{\beta_- \beta_+}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta_- \leq 0 & \forall i \in I_- \\ & \omega^\top x_i + \beta_+ \geq 0 & \forall i \in I_+ \\ & \beta_+ \leq \beta_-, \omega \in \Omega & \end{array}$$

$$\begin{aligned} \Delta &= \frac{\beta_- - \beta_+}{2} \\ \beta &= \frac{\beta_- + \beta_+}{2} \end{aligned}$$

$$\begin{array}{lll} \max & \frac{\Delta}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta & \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta & \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 & \end{array}$$

$$W(\mathcal{B}(\omega,\beta_-,\beta_+)) = \frac{\beta_-\beta_+}{\|\omega\|^\circ}$$

$$\begin{array}{ll} \max & W(\mathcal{B}(\omega,\beta_-,\beta_+)) \\ \text{s.t.} & \{x_i : i \in I_-\} \subset H(\omega,\beta_-)^- \\ & \{x_i : i \in I_+\} \subset H(\omega,\beta_+)^+ \\ & \omega \in \Omega \end{array}$$

$$\begin{array}{lll} \max & \frac{\beta_-\beta_+}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta_- \leq 0 & \forall i \in I_- \\ & \omega^\top x_i + \beta_+ \geq 0 & \forall i \in I_+ \\ & \beta_+ \leq \beta_-, \omega \in \Omega & \end{array}$$

$$\begin{array}{lll} \Delta & = & \frac{\beta_- - \beta_+}{2} \\ \beta & = & \frac{\beta_- + \beta_+}{2} \end{array}$$

$$\begin{array}{lll} \max & \frac{\Delta}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta & \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta & \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 & \end{array}$$

$$W(\mathcal{B}(\omega,\beta_-,\beta_+)) = \frac{\beta_-\beta_+}{\|\omega\|^\circ}$$

$$\begin{array}{ll} \max & W(\mathcal{B}(\omega,\beta_-,\beta_+)) \\ \text{s.t.} & \{x_i : i \in I_-\} \subset H(\omega,\beta_-)^- \\ & \{x_i : i \in I_+\} \subset H(\omega,\beta_+)^+ \\ & \omega \in \Omega \end{array}$$

$$\begin{array}{lll} \max & \frac{\beta_-\beta_+}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta_- \leq 0 & \forall i \in I_- \\ & \omega^\top x_i + \beta_+ \geq 0 & \forall i \in I_+ \\ & \beta_+ \leq \beta_-, \omega \in \Omega & \end{array}$$

$$\begin{array}{lll} \Delta & = & \dfrac{\beta_- - \beta_+}{2} \\ \beta & = & \dfrac{\beta_- + \beta_+}{2} \end{array}$$

$$\begin{array}{lll} \max & \frac{\Delta}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta & \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta & \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 & \end{array}$$

$$W(\mathcal{B}(\omega,\beta_-,\beta_+)) = \frac{\beta_-\beta_+}{\|\omega\|^\circ}$$

$$\begin{array}{ll} \max & W(\mathcal{B}(\omega,\beta_-,\beta_+)) \\ \text{s.t.} & \{x_i : i \in I_-\} \subset H(\omega,\beta_-)^- \\ & \{x_i : i \in I_+\} \subset H(\omega,\beta_+)^+ \\ & \omega \in \Omega \end{array}$$

$$\begin{array}{lll} \max & \frac{\beta_-\beta_+}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta_- \leq 0 & \forall i \in I_- \\ & \omega^\top x_i + \beta_+ \geq 0 & \forall i \in I_+ \\ & \beta_+ \leq \beta_-, \omega \in \Omega & \end{array}$$

$$\begin{array}{rcl} \Delta & = & \dfrac{\beta_- - \beta_+}{2} \\ \beta & = & \dfrac{\beta_- + \beta_+}{2} \end{array}$$

$$\begin{array}{lll} \max & \frac{\Delta}{\|\omega\|^\circ} & \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta & \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta & \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 & \end{array}$$

$$\begin{array}{ll} \max & \frac{\Delta}{\|\omega\|^\circ} \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta \quad \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 \end{array}$$

$$\begin{array}{ll} \max & \frac{1}{\|\omega\|^\circ} \\ \text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\ & \omega \in \Omega \end{array}$$

$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\ & \omega \in \Omega \end{array} \qquad (P)$$

$$\begin{array}{ll} \max & \frac{\Delta}{\|\omega\|^\circ} \\ \text{s.t.} & \omega^\top x_i + \beta \leq -\Delta \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq \Delta \quad \forall i \in I_+ \\ & \omega \in \Omega, \Delta \geq 0 \end{array}$$

$$\begin{array}{ll} \max & \frac{1}{\|\omega\|^\circ} \\ \text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\ & \omega \in \Omega \end{array}$$

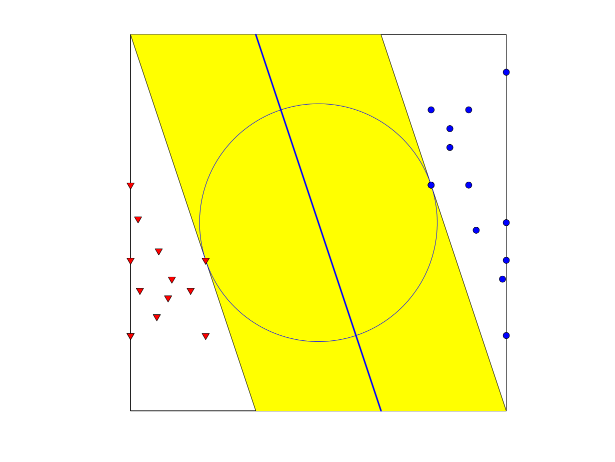
$$\begin{array}{ll} \min & \|\omega\|^\circ \\ \text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\ & \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\ & \omega \in \Omega \end{array} \qquad (P)$$

$$
\begin{array}{ll}
\max & \frac{\Delta}{\|\omega\|^\circ} \\
\text{s.t.} & \omega^\top x_i + \beta \leq -\Delta \quad \forall i \in I_- \\
& \omega^\top x_i + \beta \geq \Delta \quad \forall i \in I_+ \\
& \omega \in \Omega, \Delta \geq 0
\end{array}
$$

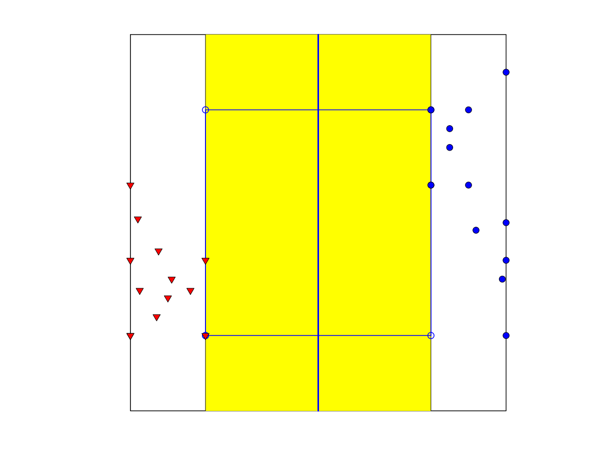
$$
\begin{array}{ll}
\max & \frac{1}{\|\omega\|^\circ} \\
\text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\
& \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\
& \omega \in \Omega
\end{array}
$$

$$
\begin{array}{ll}
\min & \|\omega\|^\circ \\
\text{s.t.} & \omega^\top x_i + \beta \leq -1 \quad \forall i \in I_- \\
& \omega^\top x_i + \beta \geq 1 \quad \forall i \in I_+ \\
& \omega \in \Omega
\end{array}
\qquad (P)
$$

Maximal width band



Maximal width band



Existence & uniqueness

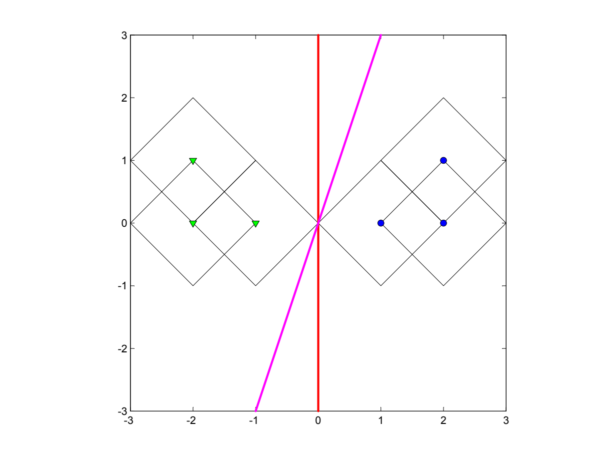
$$\begin{array}{ll} \min & \|\omega\|^{\circ} \\ \text{s.a.} & y_i(\omega^{\top} u_i + \beta) \geq 1 \, \forall i \in I \\ & \omega \in \Omega \end{array} \qquad (P)$$

Assumed to be feasible:

$$(conv(\{u_i \, : \, i \in I_+\}) + cone(\{q_j \, : \, j \in J\})) \cap conv(\{u_i \, : \, i \in I_-\}) = \emptyset$$

Th: Set $\mathcal{S}(P)$ of optimal solutions: non-empty. If $\| \cdot \|$: smooth, then $|\mathcal{S}(P)| = 1$

Uniqueness: ?



Optimality conditions

Th: Let (ω^*, β^*) : feasible for (P). Following statements are equivalent:

1. $(\omega^*, \beta^*) \in \mathcal{S}(P)$.
2. $\exists \lambda \in \mathbb{R}^I$, $\mu \in \mathbb{R}^J$ s.t.

$$
\begin{array}{rcll}
\lambda, \mu & \geq & 0 & \\
\sum_{i \in I} \lambda_i y_i & = & 0 & \\
\sum_{i \in I} \lambda_i y_i x_i + \sum_{j \in J} \mu_j q_j & \in & \partial \|\omega^*\|^\circ & \\
\lambda_i \left(y_i \left({\omega^*}^\top x_i + \beta^* \right) - 1 \right) & = & 0 & \forall i \in I \\
\mu_j \left({\omega^*}^\top q_j \right) & = & 0 & \forall j \in J
\end{array}
$$

Optimality conditions (Euclidean norm)

Th: Let (ω^*, β^*) : feasible for (P) and $\|\cdot\|$: Euclidean norm.
Following statements are equivalent:

1. $(\omega^*, \beta^*) \in \mathcal{S}(P)$.
2. $\exists \lambda \in \mathbb{R}^I$, $\mu \in \mathbb{R}^J$ s.t.

$$\begin{array}{rcll} \lambda, \mu & \geq & 0 & \\ \sum_{i \in I} \lambda_i y_i & = & 0 & \\ \sum_{i \in I} \lambda_i y_i x_i + \sum_{j \in J} \mu_j q_j & = & \omega^* & \\ \lambda_i \left(y_i \left({\omega^*}^\top x_i + \beta^* \right) - 1 \right) & = & 0 & \forall i \in I \\ \mu_j \left({\omega^*}^\top q_j \right) & = & 0 & \forall j \in J \end{array}$$

Support vectors

$$I_+(\lambda) = \{i \in I_+ : \lambda_i > 0\} \quad I_-(\lambda) = \{i \in I_- : \lambda_i > 0\}$$

$$\begin{array}{ll} I(\lambda) = I_+(\lambda) \cup I_-(\lambda) & \{x_i : \in I(\lambda)\} : \text{ support vectors at } \lambda \\ J(\mu) = \{j \in J : \mu_j > 0\} & \{q_j : j \in J(\mu)\} : \text{ support normals at } \mu \end{array}$$

Th: Let $(\omega^*, \beta^*) \in \mathcal{S}(P)$. For each $\lambda \in \mathbb{R}^I$, $\mu \in \mathbb{R}^J$, multipliers in (ω^*, β^*),

1. $I_+(\lambda) \neq \emptyset$, $I_-(\lambda) \neq \emptyset$
2. (ω^*, β^*) : also optimal for

$$\begin{array}{lll} \min & \|\omega\|^\circ & \\ \text{s.t.} & y_i(\omega^\top x_i + \beta) & \geq 1 \quad \forall i \in I(\lambda) \\ & \omega^\top q_j & \geq 0 \quad \forall j \in J(\mu). \end{array}$$

Support vectors

Th: Given $(\omega^*, \beta^*) \in \mathcal{S}(P)$, $\nexists \beta \neq \beta^*$ s.t. $(\omega^*, \beta) \in \mathcal{S}(P)$.

Th: Let $(\omega^*, \beta^*) \in \mathcal{S}(P)$. $\exists \lambda \in \mathbb{R}^I$, $\mu \in \mathbb{R}^J$, multipliers at (ω^*, β^*), s.t. $|I(\lambda) \cup J(\mu)| \leq N + 1$

Th: Let $(\omega^*, \beta^*) \in \mathcal{S}^*$, and let (λ^*, μ^*) : multiplier at (ω^*, β^*).
Given (ω, β), feasible for (P), following statements are equivalent:

1. $(\omega, \beta) \in \mathcal{S}(P)$
2. (ω, β) satisfies

$$\begin{array}{rcl} \sum_{i \in I} \lambda_i^* y_i x_i + \sum_{j \in J} \mu_j^* q_j & \in & \partial \|\omega\|^\circ \\ y_i \left(\omega^\top x_i + \beta\right) & = & 1 \;\; \forall i \in I(\lambda^*) \\ \omega^\top q_j & = & 0 \; \forall j \in J(\mu^*) \end{array}$$

3. (ω, β) satisfies:

$$\begin{array}{rcl} \omega & \in & N_B(\sum_{i \in I} \lambda_i^* y_i x_i + \sum_{j \in J} \mu_j^* q_j) \\ y_i \left(\omega^\top x_i + \beta\right) & = & 1 \;\; \forall i \in I(\lambda^*) \\ \omega^\top q_j & = & 0 \; \forall j \in J(\mu^*) \end{array}$$

$$
\begin{aligned}
\|u\|_\Omega &:= \max\left\{u^\top x : \|x\|^\circ \leq 1,\ x \in \Omega\right\}\\
&= \min\left\{\|u+q\| : q \in cone(\{q_j : j \in J\})\right\}
\end{aligned}
$$

Dual

$$
\begin{array}{ll}
\min & \|\sum_{i\in I} \mu_i y_i u_i\|_\Omega \\
\text{s.t.} & \sum_{i\in I} \mu_i y_i = 0 \\
& \sum_{i\in I} \mu_i = 1 \\
& \mu \geq 0.
\end{array}
\qquad (D)
$$

- Optimal values of (P), (D): coincide
- A maximal margin hyperplane can be obtained from optimal dual solution

$$
\begin{aligned}
\|u\|_\Omega &:= \max\left\{u^\top x : \|x\|^\circ \le 1,\, x \in \Omega\right\} \\
&= \min\left\{\|u+q\| : q \in cone(\{q_j : j \in J\})\right\}
\end{aligned}
$$

Dual

$$
\begin{array}{ll}
\min & \|\sum_{i\in I} \mu_i y_i u_i\|_\Omega \\
\text{s.t.} & \sum_{i\in I} \mu_i y_i = 0 \\
& \sum_{i\in I} \mu_i = 1 \\
& \mu \ge 0.
\end{array}
\qquad (D)
$$

- Optimal values of (P), (D): coincide
- A maximal margin hyperplane can be obtained from optimal dual solution

$$\begin{aligned} \|u\|_\Omega &:= \max\left\{u^\top x : \|x\|^\circ \leq 1, x \in \Omega\right\} \\ &= \min\left\{\|u+q\| : q \in cone(\{q_j : j \in J\})\right\} \end{aligned}$$

Dual

$$\begin{array}{ll} \min & \|\sum_{i\in I}\mu_i y_i u_i\|_\Omega \\ \text{s.t.} & \sum_{i\in I}\mu_i y_i = 0 \\ & \sum_{i\in I}\mu_i = 1 \\ & \mu \geq 0. \end{array} \qquad (D)$$

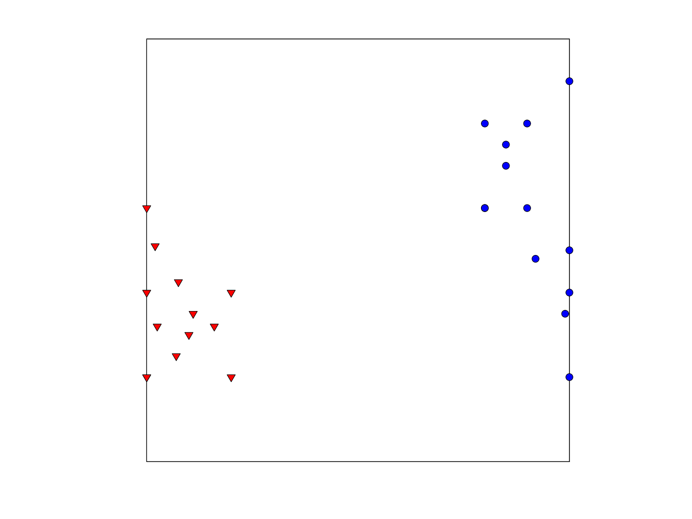
- Optimal values of (P), (D): coincide
- A maximal margin hyperplane can be obtained from optimal dual solution

Duals & closest pairs

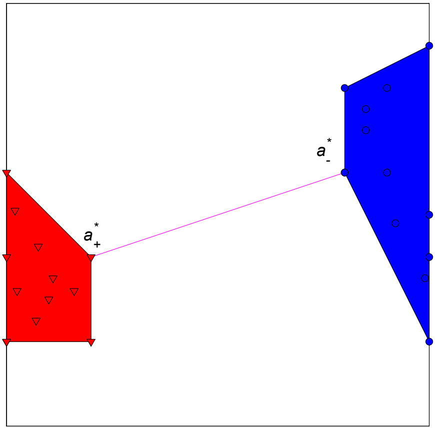
$$
\begin{array}{ll}
\min & \|a_+ + q - a_-\| \\
\text{s.a.} & a_+ \in conv\left(\{u_i : i \in I_+\}\right) \\
& a_- \in conv\left(\{u_i : i \in I_-\}\right) \\
& q \in cone\left(\{q_j : j \in J\}\right).
\end{array}
$$

- Euclidean unconstrained case:
  - $q^* = 0$
  - $a_+^* - a_-^*$: direction of maximal margin hyperplane
- In general, $a_+^* - a_-^*$: subgradient of x^*, of optimal direction

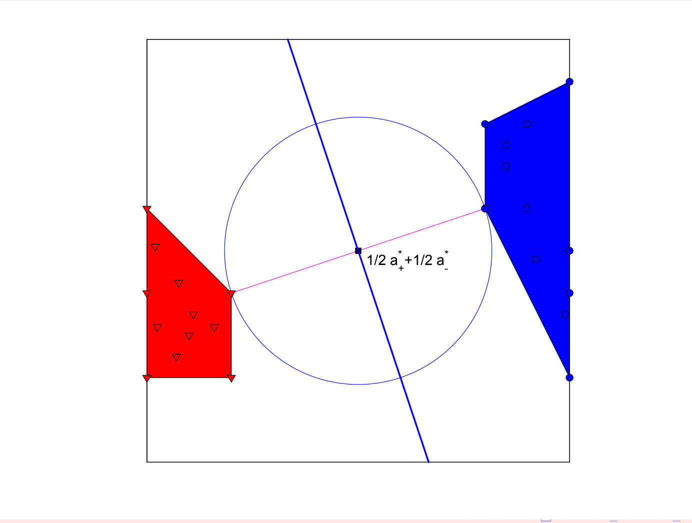
Dual & closest pairs



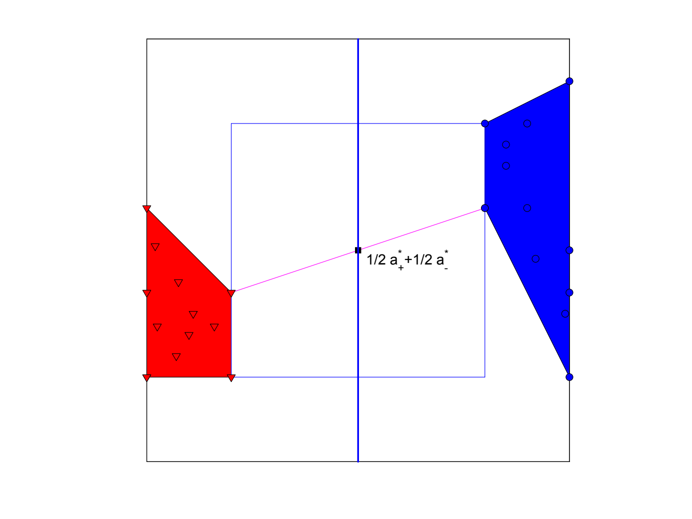
Dual & closest pairs



Dual & closest pairs



Dual & closest pairs



Two margins at same time

Margin of class c

$$\rho^c(\omega,\beta) = \min_{i\in I_c} \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}$$

Problem:

Simultaneous maximization of both margins

$$\max\{\rho^1(\omega,\beta),\ \rho^{-1}(\omega,\beta)\}$$

Two margins at same time

Margin of class c

$$\rho^c(\omega,\beta) = \min_{i\in I_c} \frac{y_i(\omega^\top x_i + \beta)}{\|\omega\|^\circ}$$

Problem:

Simultaneous maximization of both margins

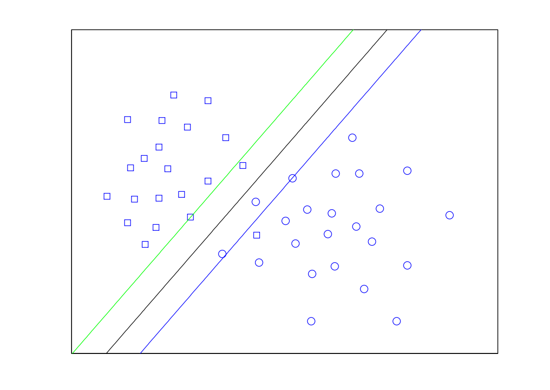
$$\max\{\rho^1(\omega,\beta),\ \rho^{-1}(\omega,\beta)\}$$

Key result

For $\|\cdot\|$: smooth, all efficient solutions anterior correspond to parallel hyperplanes.

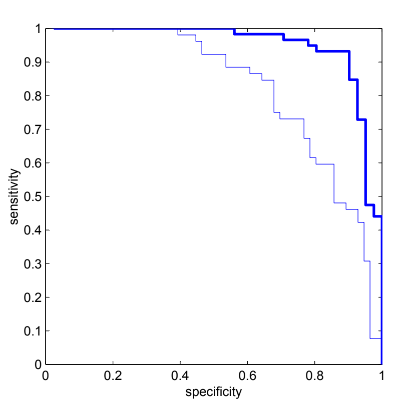
Key result

For $\|\cdot\|$: smooth, all efficient solutions anterior correspond to parallel hyperplanes.



Choice of β.

ROC curve



Non-separable populations

All the above, valid just when $\{x_i : i \in I_+\}, \{x_i : i \in I_-\}$: Ω-separable. When not the case, (P) : unfeasible, (D) : unbounded.

Possible strategies

1. change the objective function into something not requiring separability
2. modify the data, coming up with separable groups:
    - *Value embedding*
    - *Label embedding: soft margin*
3. (of course) all the above together

Non-separable populations

All the above, valid just when $\{x_i : i \in I_+\}, \{x_i : i \in I_-\}$: Ω-separable. When not the case, (P) : unfeasible, (D) : unbounded.

Possible strategies

1. change the objective function into something not requiring separability
2. modify the data, coming up with separable groups:
    - *Value embedding*
    - *Label embedding*: soft margin
3. (of course) all the above together

Non-separable populations

All the above, valid just when $\{x_i : i \in I_+\}, \{x_i : i \in I_-\}$: Ω-separable. When not the case, (P) : unfeasible, (D) : unbounded.

Possible strategies

1. change the objective function into something not requiring separability
2. modify the data, coming up with separable groups:
   - *Value embedding*
   - *Label embedding*: soft margin
3. (of course) all the above together

Non-separable populations

All the above, valid just when $\{x_i : i \in I_+\}, \{x_i : i \in I_-\}$: Ω-separable. When not the case, (P) : unfeasible, (D) : unbounded.

Possible strategies

1. change the objective function into something not requiring separability
2. modify the data, coming up with separable groups:
    - *Value embedding*
    - *Label embedding*: soft margin
3. (of course) all the above together

Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \beta \in \mathbb{R})$$

Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

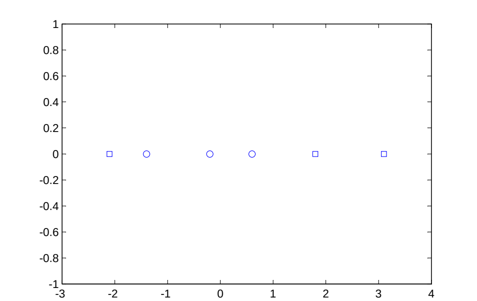
$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \; \beta \in \mathbb{R})$$

Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \ \beta \in \mathbb{R})$$

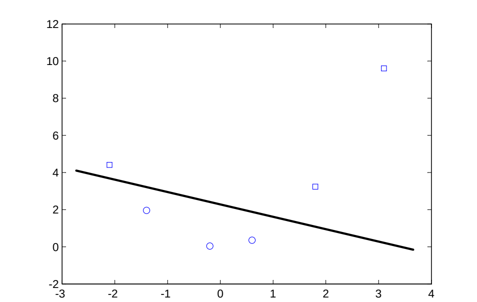


Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \; \beta \in \mathbb{R})$$



Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \, \beta \in \mathbb{R})$$

Example I

$$\begin{aligned} \varphi(s_1, s_2, \ldots, s_N) \quad = \quad & (s_1, s_2, \ldots, s_N, \\ & s_1^2, s_1 s_2, \ldots, s_1 s_N, \\ & s_2^2, \ldots, s_2 s_N, \\ & \ldots s_N^2) \end{aligned}$$

"linear" classifier of form $x^\top W x + \beta$

Value embedding

Take $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^T$, and replace each (x_i, y_i) by $(\varphi(x_i), y_i)$

"Linear" classifier sought:

$$\omega^\top \varphi(x) + \beta \qquad (\omega \in \mathbb{R}^T \; \beta \in \mathbb{R})$$

Example II (patterns)

$$\begin{aligned} \tau_{i1} > \tau_{i2} > \ldots > \tau_{ir} \quad &: \quad \text{given } \forall i \\ \varphi(s_1, s_2, \ldots, s_N) &= (\delta(s_i - \tau_{ik}))_{ik} \\ \delta(s_i - \tau_{ik}) &= \begin{cases} 1, & \text{if } s_i - \tau_{ik} > 0 \\ 0, & \text{else} \end{cases} \end{aligned}$$

Combining the two above,

database	$\|\mathcal{O}\|$	N
sonar	208	60
ionosphere	351	34
wdbc	569	30

CART		
database	training	testing
sonar	0.98	0.77
ionosphere	0.98	0.88
wdbc	0.99	0.94

SVM		
database	training	testing
sonar	1.00	0.84
ionosphere	1.00	0.92
wdbc	1.00	0.96

Combining the two above,

database	$\lvert\mathcal{O}\rvert$	N
sonar	208	60
ionosphere	351	34
wdbc	569	30

CART		
database	training	testing
sonar	0.98	0.77
ionosphere	0.98	0.88
wdbc	0.99	0.94

SVM		
database	training	testing
sonar	1.00	0.84
ionosphere	1.00	0.92
wdbc	1.00	0.96

Combining the two above,

database	$\|\mathcal{O}\|$	N
sonar	208	60
ionosphere	351	34
wdbc	569	30

CART		
database	training	testing
sonar	0.98	0.77
ionosphere	0.98	0.88
wdbc	0.99	0.94

SVM		
database	training	testing
sonar	1.00	0.84
ionosphere	1.00	0.92
wdbc	1.00	0.96

Kernels (Euclidean norm)

$$\begin{array}{ll} \max & \sum_{i\in I}\lambda_i - \frac{1}{2}\|\sum_i \lambda_i y_i x_i\|^2 \\ \text{s.t.} & \sum_{i\in I}\lambda_i y_i = 0, \quad \lambda \geq 0 \end{array}$$

$$(x_i, y_i) \xrightarrow{\varphi} (\varphi(x_i), y_i)$$

$$\begin{array}{ll} \max & \sum_{i\in I}\lambda_i - \frac{1}{2}\sum_{i,j\in I}\lambda_i\lambda_j y_i y_j \varphi(x_i)^\top \varphi(x_j) \\ \text{s.t.} & \sum_{i\in I}\lambda_i y_i = 0, \quad \lambda \geq 0 \end{array}$$

$$\omega = \sum_i \lambda_i y_i \varphi(x_i)$$

Classification rule:

$$\begin{array}{lll} \text{assign } x \text{ to group 1} & \text{iff} & \omega^\top \varphi(x) + \beta > 0 \\ & \text{iff} & \sum_i \lambda_i y_i \varphi(x_i)^\top \varphi(x) + \beta > 0 \end{array}$$

Kernels (Euclidean norm)

$$\begin{array}{ll} \max & \sum_{i\in I}\lambda_i - \frac{1}{2}\|\sum_i \lambda_i y_i x_i\|^2 \\ \text{s.t.} & \sum_{i\in I}\lambda_i y_i = 0, \quad \lambda \geq 0 \end{array}$$

$$(x_i, y_i) \xrightarrow{\varphi} (\varphi(x_i), y_i)$$

$$\begin{array}{ll} \max & \sum_{i\in I}\lambda_i - \frac{1}{2}\sum_{i,j\in I}\lambda_i\lambda_j y_i y_j K(x_i, x_j) \\ \text{s.t.} & \sum_{i\in I}\lambda_i y_i = 0, \quad \lambda \geq 0 \end{array}$$

$$\omega = \sum_i \lambda_i y_i \varphi(x_i)$$

Classification rule:

$$\begin{array}{lll} \text{assign } x \text{ to group 1} & \text{iff} & \omega^\top \varphi(x) + \beta > 0 \\ & \text{iff} & \sum_i \lambda_i y_i K(x_i, x) + \beta > 0 \end{array}$$

Soft margin. Cortes-Vapnik

Replaces (P) by one in the form

$$\begin{array}{ll} \min & \|\omega\|_2^2 + C\sum_{i\in I}\eta_i^p \\ \text{s.t.} & y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i \,\forall i \in I \\ & \eta \geq 0 \end{array}$$

for $p \geq 1$ (typically $p = 1$) and $C > 0$, or

$$\begin{array}{ll} \min & \|\omega\|_1 + C\sum_{i\in I}\eta_i \\ \text{s.t.} & y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i \,\forall i \in I \\ & \eta \geq 0. \end{array}$$

Value embedding

- Soft margin: almost particular cases
- enables us to use directly results of the separable case

$$(x_i, y_i) \longrightarrow (\widehat{x}_i, y_i) := (\overbrace{(x_i; 0, \overset{(i)}{y_i}, 0, \ldots, 0)}^{\widehat{x}}, y_i) \in (\mathbb{R}^N \times \mathbb{R}^I) \times \{-1, 1\}$$

Sets $\{\widehat{x}_i : i \in I_+\}$ y $\{\widehat{x}_i : i \in I_-\}$: separable.

Any method applicable to the separable case can be used to the non-separable case after embedding:

Value embedding

- Soft margin: almost particular cases
- enables us to use directly results of the separable case

$$(x_i, y_i) \longrightarrow (\widehat{x}_i, y_i) := (\overbrace{(x_i; 0, \overset{(i)}{y_i}, 0, \ldots, 0)}^{\widehat{x}}, y_i) \in (\mathbb{R}^N \times \mathbb{R}^I) \times \{-1, 1\}$$

Sets $\{\widehat{x}_i : i \in I_+\}$ y $\{\widehat{x}_i : i \in I_-\}$: separable.

Any method applicable to the separable case can be used to the non-separable case after embedding:

Value embedding

- Soft margin: almost particular cases
- enables us to use directly results of the separable case

$$(x_i, y_i) \longrightarrow (\widehat{x_i}, y_i) := (\overbrace{(x_i; 0, \overset{(i)}{y_i}, 0, \ldots, 0)}^{\widehat{x}}, y_i) \in (\mathbb{R}^N \times \mathbb{R}^I) \times \{-1, 1\}$$

Sets $\{\widehat{x_i} : i \in I_+\}$ y $\{\widehat{x_i} : i \in I_-\}$: separable.

Any method applicable to the separable case can be used to the non-separable case after embedding:

$$\begin{array}{ll} \min & \|\widehat{\omega}\|^\circ \\ \text{s.t.} & y_i\left(\widehat{\omega}^\top \widehat{x_i} + \beta\right) \geq 1 \quad \forall i \in I \end{array}$$

Value embedding

- Soft margin: almost particular cases
- enables us to use directly results of the separable case

$$(x_i, y_i) \longrightarrow (\widehat{x}_i, y_i) := (\overbrace{(x_i; 0, \overset{(i)}{y_i}, 0, \ldots, 0)}^{\widehat{x}}, y_i) \in (\mathbb{R}^N \times \mathbb{R}^I) \times \{-1, 1\}$$

Sets $\{\widehat{x}_i : i \in I_+\}$ y $\{\widehat{x}_i : i \in I_-\}$: separable.

Any method applicable to the separable case can be used to the non-separable case after embedding:

$$\begin{array}{ll} \min & \|\widehat{\omega}\|^\circ \\ \text{s.t.} & y_i \left(\widehat{\omega}^\top \widehat{x}_i + \beta\right) \geq 1 \quad \forall i \in I \end{array}$$

Value embedding

- Soft margin: almost particular cases
- enables us to use directly results of the separable case

$$(x_i, y_i) \longrightarrow (\widehat{x_i}, y_i) := (\overbrace{(x_i; 0, \overset{(i)}{y_i}, 0, \dots, 0)}^{\widehat{x}}, y_i) \in (\mathbb{R}^N \times \mathbb{R}^I) \times \{-1, 1\}$$

Sets $\{\widehat{x_i} : i \in I_+\}$ y $\{\widehat{x_i} : i \in I_-\}$: separable.

Any method applicable to the separable case can be used to the non-separable case after embedding:

$$\begin{array}{ll} \min & \|\widehat{\omega}\|^\circ \\ \text{s.t.} & y_i \left(\widehat{\omega}^\top \widehat{x_i} + \beta\right) \geq 1 \quad \forall i \in I \end{array}$$

Assumptions

$$\|(\omega, \eta)\| = \gamma(\gamma_1(\omega), \gamma_2(\eta)),$$

- γ_1 : norm in $\mathbb{R}^N$
- γ_2 : norm in $\mathbb{R}^I$, monotonic in $\mathbb{R}^I_+$
- $\gamma(t_1, t_2) = \max\{|t_1|, \frac{1}{C}|t_2|\}$, thus $\gamma^\circ(t_1, t_2) = |t_1| + C|t_2|$

Assumptions

$$\|(\omega,\eta)\| = \gamma(\gamma_1(\omega),\gamma_2(\eta)),$$

- γ_1 : norm in $\mathbb{R}^N$
- γ_2 : norm in $\mathbb{R}^I$, monotonic in $\mathbb{R}^I_+$
- $\gamma(t_1,t_2) = \max\{|t_1|, \frac{1}{C}|t_2|\}$, thus $\gamma^\circ(t_1,t_2) = |t_1| + C|t_2|$

Primal (P_{soft})	
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$
s.t.	$y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i \quad \forall i \in I$

Assumptions

$$\|(\omega,\eta)\| = \gamma(\gamma_1(\omega),\gamma_2(\eta)),$$

- γ_1 : norm in $\mathbb{R}^N$
- γ_2 : norm in $\mathbb{R}^I$, monotonic in $\mathbb{R}^I_+$
- $\gamma(t_1,t_2) = \max\{|t_1|, \frac{1}{C}|t_2|\}$, thus $\gamma^\circ(t_1,t_2) = |t_1| + C|t_2|$

Primal (P_{soft})	
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ\left(\left(1 - y_i\left(\omega^\top x_i + \beta\right)\right)^+_{i\in I}\right)$
s.t.	no constraints!

Assumptions

$$\|(\omega,\eta)\| = \gamma(\gamma_1(\omega),\gamma_2(\eta)),$$

- γ_1 : norm in $\mathbb{R}^N$
- γ_2 : norm in $\mathbb{R}^I$, monotonic in $\mathbb{R}^I_+$
- $\gamma(t_1,t_2) = \max\{|t_1|, \frac{1}{C}|t_2|\}$, thus $\gamma^\circ(t_1,t_2) = |t_1| + C|t_2|$

Dual (D_{soft})	
min	$\max(\gamma_1(\sum_{i\in I}\mu_i y_i x_i), C\gamma_2(\mu))$
s.t.	$\sum_{i\in I}\mu_i y_i = 0$
	$\sum_{i\in I}\mu_i = 1$
	$\mu \geq 0$

Primal (P_{soft})	
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$
s.t.	$y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i \quad \forall i \in I$

Proposition

Let $(\omega^*, \eta^*, \beta^*)$, feasible for (P_{soft}). Following statements are equivalent:

1. $(\omega^*, \eta^*, \beta^*)$: optimal for (P_{soft})
2. $\exists \lambda \in \mathbb{R}^I$ s.t.

$$\begin{cases} \lambda & \geq & 0 \\ \sum_{i\in I} \lambda_i y_i & = & 0 \\ \sum_{i\in I} \lambda_i y_i x_i & \in & \partial\gamma_1^\circ(\omega^*) \\ \lambda & \in & C\partial\gamma_2^\circ(\eta^*) \\ \lambda_i\left(y_i\left({\omega^*}^\top x_i + \beta^*\right) - 1 + \eta_i^*\right) & = & 0 \quad \forall i \in I \end{cases}$$

Primal (P_{soft})		
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$	
s.t.	$y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i$	$\forall i \in I$

Existence: ok

Proposition

Assume γ_1 is smooth. Let $(\omega^*, \eta^*, \beta^*)$: optimal to (P_{soft}). Any optimal solution to (ω, η, β) satisfies

$$\omega = \nu\omega^*$$

for some $\nu > 0$. In other words, all optimal solutions to (P_{soft}) generate parallel hyperplanes in $\mathbb{R}^N$.

Primal (P_{soft})		
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$	
s.t.	$y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i$	$\forall i \in I$

Existence: ok

Proposition

Assume γ_1 is smooth. Let $(\omega^*, \eta^*, \beta^*)$: optimal to (P_{soft}). Any optimal solution to (ω, η, β) satisfies

$$\omega = \nu\omega^*$$

for some $\nu > 0$. In other words, all optimal solutions to (P_{soft}) generate parallel hyperplanes in $\mathbb{R}^N$.

Degenerate case: $\omega^* = 0, \eta^* \neq 0, \beta^* = 0$

Proposition

Following statements are equivalent:

1. $\exists \eta^* \in \mathbb{R}^I$ s.t. $(0, \eta^*, 0)$: optimal to (P_{soft})
2. $\exists \lambda \in \mathbb{R}^I$ s.t.

$$
\begin{array}{rcl}
\lambda & \geq & 0 \\
\sum_{i \in I} \lambda_i y_i & = & 0 \\
\lambda & \in & \partial \gamma_2^\circ(e) \\
\gamma_1(\sum_{i \in I} \lambda_i y_i x_i &) \leq & C
\end{array}
$$

Degenerate case: $\omega^* = 0, \eta^* \neq 0, \beta^* = 0$

Primal (P_{soft})		
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$	
s.t.	$y_i \left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i$	$\forall i \in I$

Proposition

Following statements are equivalent:

1. $\forall C > 0 \ \exists$ optimal solution to (P_{soft}) in form $(0, \eta, 0)$
2. $\exists \lambda \in \mathbb{R}^I$ s.t.

$$\begin{aligned} \lambda &\geq 0 \\ \sum_{i \in I} \lambda_i y_i &= 0 \\ \lambda &\in \partial\gamma_2^\circ(e) \\ \sum_{i \in I} \lambda_i y_i x_i &= 0 \end{aligned}$$

Primal (P_{soft})		
min	$\gamma_1^\circ(\omega) + C\gamma_2^\circ(\eta)$	
s.t.	$y_i\left(\omega^\top x_i + \beta\right) \geq 1 - \eta_i$	$\forall i \in I$

Proposition

Assume γ_2 : symmetrical and γ_2° : differentiable at e. Following statements are equivalent:

1. For all $C > 0$, optimal to (P_{soft}) has form $(0, \eta, 0)$
2. $\sum_{i\in I} y_i = 0$, $\sum_{i\in I} y_i x_i = 0$

In practice we can easily avoid trivial solution $\omega = 0, \beta = 0$: drop just one element from database! (with this, $\sum_{i\in I} y_i \neq 0$)

Time for commercials

Dirección http://www.euro2006.org/

EURO XXI in Iceland
July 2-5, 2006

21st European Conference on Operational Research

- Home
- Organisation
- General Information
- Registration
- Call for Papers
- Abstract Submission
- Program Schedule
- Awards
- Conference Venue
- Exhibition
- Sponsors
- Social Events
- Travel
- Accommodation
- Iceland

List of currently accepted streams (still preliminary):

Stream	Organizer
Adaptive Memory Programming	Cesar Rego (crego@bus.olemiss.edu)
Behavioural and Experimental Economics	Bernd Brandl (bernd.brandl@univie.ac.at) Johannes Leitner (johannesleitner@gmx.de)
Business Sessions	Bjarni Kristjánsson (bjarni@maximalsoftware.com)
Combinatorial Optimization	Silvano Martello (smartello@deis.unibo.it)
Complex Societal Problems	Dorien DeTombe (DeTombe@lri.jur.uva.nl)
Computational Biology and Medicine	Jacek Blazewicz (Jacek.Blazewicz@cs.put.poznan.pl) Metin Türkay (mturkay@ku.edu.tr)
Convex Optimization Methods	Marc Teboulle (teboulle@post.tau.ac.il)
Cutting and Packing	José Fernando Oliveira (jfo@fe.up.pt)
Data Mining	Emilio Carrizosa (ecarrizosa@us.es) Jacob Kogan (kogan@math.umbc.edu)

Decision tables

$\mathcal{D}$	Crit_1	Crit_2	$\ldots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\ldots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\ldots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\ldots$	$\psi_N(d_m)$

Aim

Associate with criteria scores $\omega_1, \omega_2, \ldots, \omega_N$ reflecting decision-maker's preferences.

Possible approaches

- Direct scoring
- Pairwise comparison of criteria
- Pairwise comparison of actions

Decision tables

$\mathcal{D}$	Crit_1	Crit_2	$\ldots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\ldots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\ldots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\ldots$	$\psi_N(d_m)$

Aim

Associate with criteria scores $\omega_1, \omega_2, \ldots, \omega_N$ reflecting decision-maker's preferences.

Possible approaches

- Direct scoring
- Pairwise comparison of criteria
- Pairwise comparison of actions

Decision tables

$\mathcal{D}$	Crit_1	Crit_2	$\ldots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\ldots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\ldots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\ldots$	$\psi_N(d_m)$

Aim

Associate with criteria scores $\omega_1, \omega_2, \ldots, \omega_N$ reflecting decision-maker's preferences.

Possible approaches

- Direct scoring
- Pairwise comparison of criteria
- Pairwise comparison of actions

Decision tables

$\mathcal{D}$	Crit_1	Crit_2	$\ldots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\ldots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\ldots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\ldots$	$\psi_N(d_m)$

Aim

Associate with criteria scores $\omega_1, \omega_2, \ldots, \omega_N$ reflecting decision-maker's preferences.

Possible approaches

- Direct scoring
- Pairwise comparison of criteria
- Pairwise comparison of actions

Now, suppose given ...

- Decision table:

$\mathcal{D}$	Crit_1	Crit_2	$\dots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\dots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\dots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\dots$	$\psi_N(d_m)$

- Binary irreflexive relation $P \subset \mathcal{D} \times \mathcal{D}$

- ω sought s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D},\ dPd'$
- ω induces binary relations on $\mathcal{D}$

$$\begin{array}{llll} dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') & \text{(strict preference)} \\ dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d') & \text{(indifference)} \end{array}$$

Now, suppose given ...

- Decision table:

$\mathcal{D}$	Crit_1	Crit_2	$\dots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\dots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\dots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\dots$	$\psi_N(d_m)$

- Binary irreflexive relation $P \subset \mathcal{D} \times \mathcal{D}$

- ω sought s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D},\ dPd'$
- ω induces binary relations on $\mathcal{D}$

$$\begin{array}{lll} dP_\omega d' & \text{iff} \quad \omega^\top \Psi(d) > \omega^\top \Psi(d') & \text{(strict preference)} \\ dI_\omega d' & \text{iff} \quad \omega^\top \Psi(d) = \omega^\top \Psi(d') & \text{(indifference)} \end{array}$$

Now, suppose given ...

- Decision table:

$\mathcal{D}$	Crit_1	Crit_2	...	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	...	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	...	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	...	$\psi_N(d_m)$

- Binary irreflexive relation $P \subset \mathcal{D} \times \mathcal{D}$

- ω sought s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D},\ dPd'$
- ω induces binary relations on $\mathcal{D}$

$$\begin{array}{lll l} dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') & \text{(strict preference)} \\ dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d') & \text{(indifference)} \end{array}$$

Now, suppose given ...

- Decision table:

$\mathcal{D}$	Crit_1	Crit_2	$\dots$	Crit_N
d_1	$\psi_1(d_1)$	$\psi_2(d_1)$	$\dots$	$\psi_N(d_1)$
d_2	$\psi_1(d_2)$	$\psi_2(d_2)$	$\dots$	$\psi_N(d_2)$
$\vdots$	$\vdots$		$\ddots$	$\vdots$
d_m	$\psi_1(d_m)$	$\psi_2(d_m)$	$\dots$	$\psi_N(d_m)$

- Binary irreflexive relation $P \subset \mathcal{D} \times \mathcal{D}$

- ω sought s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D},\ dPd'$
- ω induces binary relations on $\mathcal{D}$

$$\begin{array}{lll l} dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') & \text{(strict preference)} \\ dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d') & \text{(indifference)} \end{array}$$

ω s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall dPd'$

$$
\begin{array}{rcl}
dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') \\
dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d')
\end{array}
$$

- P_ω :
  - compatible with P ($P \subset P_\omega$)
  - enjoys nicer properties
- By construction, each ω_j measures importance of Ψ_j in P_ω
- Hence ω_j can be seen as measure of importance of Ψ_j in P

ω s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall dPd'$

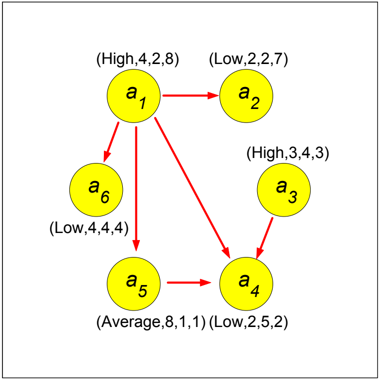
$$\begin{array}{rcl} dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') \\ dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d') \end{array}$$

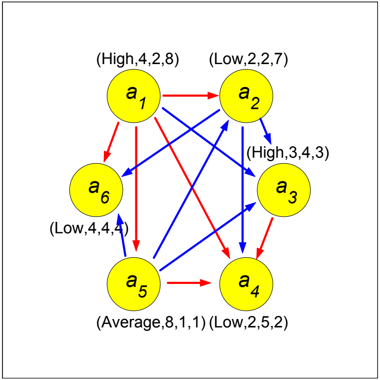
- P_ω :
  - compatible with P ($P \subset P_\omega$)
  - enjoys nicer properties
- By construction, each ω_j measures importance of Ψ_j in P_ω
- Hence ω_j can be seen as measure of importance of Ψ_j in P

ω s.t. $\omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall dPd'$

$$\begin{array}{rcl} dP_\omega d' & \text{iff} & \omega^\top \Psi(d) > \omega^\top \Psi(d') \\ dI_\omega d' & \text{iff} & \omega^\top \Psi(d) = \omega^\top \Psi(d') \end{array}$$

- P_ω :
  - compatible with P ($P \subset P_\omega$)
  - enjoys nicer properties
- By construction, each ω_j measures importance of Ψ_j in P_ω
- Hence ω_j can be seen as measure of importance of Ψ_j in P


(High,4,2,8)
(Low,2,2,7)
a_1
a_2
(High,3,4,3)
a_6
a_3
(Low,4,4,4)
a_5
a_4
(Average,8,1,1)
(Low,2,5,2)


(High,4,2,8)
(Low,2,2,7)
a_1
a_2
(High,3,4,3)
a_6
a_3
(Low,4,4,4)
a_5
a_4
(Average,8,1,1)
(Low,2,5,2)

	C_1 (max)	C_2 (max)	C_3 (max)	C_4 (max)
d_1	High	4	2	8
d_2	Low	2	2	7
d_3	High	3	4	3
d_4	High	2	5	2
d_5	Average	8	1	1
d_6	Low	4	4	4

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

$$
\begin{aligned}
\Psi_{11}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "High"} \\ 0, & \text{else} \end{cases} \\
\Psi_{12}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Average"} \\ 0, & \text{else} \end{cases} \\
\Psi_{13}(a) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Low"} \\ 0, & \text{else} \end{cases} \\
\Psi_i(d) &= \text{score of } d \text{ according to } C_i, \qquad i = 2, \ldots, 4
\end{aligned}
$$

	C_1 (max)	C_2 (max)	C_3 (max)	C_4 (max)
d_1	High	4	2	8
d_2	Low	2	2	7
d_3	High	3	4	3
d_4	High	2	5	2
d_5	Average	8	1	1
d_6	Low	4	4	4

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

$$\begin{aligned}
\Psi_{11}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "High"} \\ 0, & \text{else} \end{cases} \\
\Psi_{12}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Average"} \\ 0, & \text{else} \end{cases} \\
\Psi_{13}(a) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Low"} \\ 0, & \text{else} \end{cases} \\
\Psi_i(d) &= \text{score of } d \text{ according to } C_i, \qquad i = 2, \ldots, 4
\end{aligned}$$

	C_1 (max)	C_2 (max)	C_3 (max)	C_4 (max)
d_1	High	4	2	8
d_2	Low	2	2	7
d_3	High	3	4	3
d_4	High	2	5	2
d_5	Average	8	1	1
d_6	Low	4	4	4

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

$$
\begin{aligned}
\Psi_{11}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "High"} \\ 0, & \text{else} \end{cases} \\
\Psi_{12}(d) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Average"} \\ 0, & \text{else} \end{cases} \\
\Psi_{13}(a) &= \begin{cases} 1, & \text{if } d \text{ in } C_1 \text{ takes the value "Low"} \\ 0, & \text{else} \end{cases} \\
\Psi_i(d) &= \text{score of } d \text{ according to } C_i, \qquad i = 2, \ldots, 4
\end{aligned}
$$

Partial information

Constraints on ω

Assume given Ω : non-empty polyhedral cone

$$\Omega = \left\{\omega : \; q_j^\top \omega \geq 0 \qquad \forall j \in J\right\}$$

$(0 \leq |J| < \infty)$

- We seek ω in Ω_P,

$$\Omega_P = \left\{\omega \in \Omega : \; \omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall dPd'\right\}$$

$\Omega_P \neq \emptyset$ iff $0 \notin \left(conv(\{\Psi(d) - \Psi(d') : dPd'\}) + cone(\{q_j : j \in J\})\right)$

Partial information

Constraints on ω

Assume given Ω : non-empty polyhedral cone

$$\Omega = \left\{\omega : \; q_j^\top \omega \geq 0 \qquad \forall j \in J\right\}$$

$(0 \leq |J| < \infty)$

- We seek ω in Ω_P,

$$\Omega_P = \left\{\omega \in \Omega : \; \omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall dPd'\right\}$$

$\Omega_P \neq \emptyset$ iff $0 \notin \left(conv(\{\Psi(d) - \Psi(d') : dPd'\}) + cone(\{q_j : j \in J\})\right)$

Ω

Examples

- Criteria: "the higher, the better" thus

$$\omega_j \geq 0$$

- $\omega_i \leq K_{ij}\omega_j$
- If d, d' : known to be indifferent,

$$\begin{aligned} (\Psi(d) - \Psi(d'))^\top \omega &\geq 0 \\ (\Psi(d') - \Psi(d))^\top \omega &\geq 0 \end{aligned}$$

- ...

$\Omega_P = \left\{ x \in \Omega : \; x^\top \Psi(d) > x^\top \Psi(d') \qquad \forall dPd' \right\}$

- In case $\omega \in \Omega_P$ exists, then highly desirable that $\omega^\top \Psi(d) - \omega^\top \Psi(d')$ should be not only positive but high in all pairs $d, d' \in \mathcal{D}, dPd'$
- Achieved if we maximize the lowest slack,

$$\min_{dPd'} \left\{ \omega^\top \Psi(d) - \omega^\top \Psi(d') \right\}$$

- Maximizing over Ω_P the lowest slack is, as soon as $\Omega_P \neq \emptyset$, an optimization problem with unbounded solution
- Hence, normalization condition which enables us to identify ω and $\vartheta\omega$, $\vartheta > 0$

$$\max_{\omega \in \Omega} \frac{\min_{dPd'} \left\{ \omega^\top \Psi(d) - \omega^\top \Psi(d') \right\}}{\gamma(\omega)}$$

$\Omega_P = \left\{ x \in \Omega : \; x^\top \Psi(d) > x^\top \Psi(d') \qquad \forall dPd' \right\}$

- In case $\omega \in \Omega_P$ exists, then highly desirable that $\omega^\top \Psi(d) - \omega^\top \Psi(d')$ should be not only positive but high in all pairs $d, d' \in \mathcal{D}, dPd'$
- Achieved if we maximize the lowest slack,

$$\min_{dPd'} \left\{ \omega^\top \Psi(d) - \omega^\top \Psi(d') \right\}$$

- Maximizing over Ω_P the lowest slack is, as soon as $\Omega_P \neq \emptyset$, an optimization problem with unbounded solution
- Hence, normalization condition which enables us to identify ω and $\vartheta\omega$, $\vartheta > 0$

$$\max_{\omega \in \Omega} \frac{\min_{dPd'} \left\{ \omega^\top \Psi(d) - \omega^\top \Psi(d') \right\}}{\gamma(\omega)}$$

$\Omega_P = \left\{x \in \Omega : \ x^\top \Psi(d) > x^\top \Psi(d') \qquad \forall dPd'\right\}$

- In case $\omega \in \Omega_P$ exists, then highly desirable that $\omega^\top \Psi(d) - \omega^\top \Psi(d')$ should be not only positive but high in all pairs $d, d' \in \mathcal{D}, dPd'$
- Achieved if we maximize the lowest slack,

$$\min_{dPd'} \left\{\omega^\top \Psi(d) - \omega^\top \Psi(d')\right\}$$

- Maximizing over Ω_P the lowest slack is, as soon as $\Omega_P \neq \emptyset$, an optimization problem with unbounded solution
- Hence, normalization condition which enables us to identify ω and $\vartheta\omega$, $\vartheta > 0$

$$\max_{\omega\in\Omega} \frac{\min_{dPd'} \left\{\omega^\top \Psi(d) - \omega^\top \Psi(d')\right\}}{\gamma(\omega)}$$

$\Omega_P = \left\{x \in \Omega : \; x^\top \Psi(d) > x^\top \Psi(d') \qquad \forall dPd'\right\}$

- In case $\omega \in \Omega_P$ exists, then highly desirable that $\omega^\top \Psi(d) - \omega^\top \Psi(d')$ should be not only positive but high in all pairs $d, d' \in \mathcal{D}, dPd'$
- Achieved if we maximize the lowest slack,

$$\min_{dPd'} \left\{\omega^\top \Psi(d) - \omega^\top \Psi(d')\right\}$$

- Maximizing over Ω_P the lowest slack is, as soon as $\Omega_P \neq \emptyset$, an optimization problem with unbounded solution
- Hence, normalization condition which enables us to identify ω and $\vartheta\omega$, $\vartheta > 0$

$$\max_{\omega \in \Omega} \frac{\min_{dPd'} \left\{\omega^\top \Psi(d) - \omega^\top \Psi(d')\right\}}{\gamma(\omega)}$$

$$\max_{\omega \in \Omega} \frac{\min_{dPd'} \left\{ \omega^\top \Psi(d) - \omega^\top \Psi(d') \right\}}{\gamma(\omega)}$$

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top \Psi(d) - \omega^\top \Psi(d') \geq 1 \quad \forall dPd' \\ & \omega \in \Omega \end{array}$$

If

- $\Omega \subset \mathbb{R}^N_+$
- $\gamma : \ell_1$

then $\gamma(\omega) = e^\top \omega \, \forall \omega \in \Omega$, thus

$$\begin{array}{ll} \min & e^\top \omega \\ \text{s.t.} & \omega^\top \Psi(d) - \omega^\top \Psi(d') \geq 1 \quad \forall dPd' \\ & \omega \in \Omega \end{array}$$

$$\max_{\omega\in\Omega} \frac{\min_{dPd'}\left\{\omega^\top\Psi(d)-\omega^\top\Psi(d')\right\}}{\gamma(\omega)}$$

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top\Psi(d)-\omega^\top\Psi(d')\geq 1 \quad \forall dPd' \\ & \omega\in\Omega \end{array}$$

If

- $\Omega\subset\mathbb{R}^N_+$
- $\gamma:\ell_1$

then $\gamma(\omega)=e^\top\omega\,\forall\omega\in\Omega$, thus

$$\begin{array}{ll} \min & e^\top\omega \\ \text{s.t.} & \omega^\top\Psi(d)-\omega^\top\Psi(d')\geq 1 \quad \forall dPd' \\ & \omega\in\Omega \end{array}$$

$$\max_{\omega \in \Omega} \frac{\min_{dPd'} \left\{\omega^\top \Psi(d) - \omega^\top \Psi(d')\right\}}{\gamma(\omega)}$$

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top \Psi(d) - \omega^\top \Psi(d') \geq 1 \quad \forall dPd' \\ & \omega \in \Omega \end{array}$$

If

- $\Omega \subset \mathbb{R}_+^N$
- $\gamma : \ell_1$

then $\gamma(\omega) = e^\top \omega \, \forall \omega \in \Omega$, thus

$$\begin{array}{ll} \min & e^\top \omega \\ \text{s.t.} & \omega^\top \Psi(d) - \omega^\top \Psi(d') \geq 1 \quad \forall dPd' \\ & \omega \in \Omega \end{array}$$

Example

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top(\Psi(d) - \Psi(d')) \geq 1 \quad \forall dPd' \\ & \omega \in \Omega_P \end{array}$$

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

Example

$$
\begin{array}{ll}
\min & \gamma(\omega) \\
\text{s.t.} & \omega^\top(\Psi(d) - \Psi(d')) \geq 1 \quad \forall dPd' \\
& \omega \in \Omega_P
\end{array}
$$

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

Constraints

- $\omega \geq 0$
- $\omega_{11} \geq \omega_{12} \geq \omega_{13}$
- $\omega_4 \leq 5\omega_3$

Example

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top(\Psi(d) - \Psi(d')) \geq 1 \quad \forall dPd' \\ & \omega \in \Omega_P \end{array}$$

	Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_2	Ψ_3	Ψ_4
d_1	1	0	0	4	2	8
d_2	0	0	1	2	2	7
d_3	1	0	0	3	4	3
d_4	0	0	1	2	5	2
d_5	0	1	0	8	1	1
d_6	0	0	1	4	4	4

Solution

- $(\omega_{11}, \omega_{12}, \omega_{13}, \omega_2, \omega_3, \omega_4) = (0.3418, 0.0, 0.0, 0.3397, 0.0531, 0.2654)$
- $\omega^\top \Psi(d_i) = (3.9302, 2.6435, 2.3695, 1.4757, 3.0363, 2.6329)$
- $d_1 P_\omega d_5 P_\omega d_2 P_\omega d_6 P_\omega d_3 P_\omega d_4$

$$
\begin{array}{rcl}
I_+ & := & \{(d,d') \in \mathcal{D} \times \mathcal{D} :\ dPd'\} \\
I_- & := & \{(d',d) \in \mathcal{D} \times \mathcal{D} :\ dPd'\} \\
u_{dd'} & := & \Psi(d) - \Psi(d') \in \mathbb{R}^N \qquad \forall (d,d') \in I (I := I_+ \cup I_-) \\
\Omega_P & := & \left\{\omega \in \Omega :\ \omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D},\ dPd'\right\} \\
\|\cdot\| & := & \gamma^\circ
\end{array}
$$

(P)

$$
\begin{array}{ll}
\min & \|\omega\|^\circ \\
\text{s.t.} & y_i(\omega^\top u_i + \beta) \geq 1\, \forall i \in I \\
& \omega \in \Omega
\end{array}
$$

(Q)

$$
\begin{array}{ll}
\min & \gamma(\omega) \\
\text{s.t.} & y_{dd'}(\omega^\top u_{dd'} + \beta) \geq 1\, \forall (d,d') \in I \\
& \omega \in \Omega
\end{array}
$$

Th: If (ω, β) : optimal to (Q), then $\beta = 0$

$$
\begin{array}{rcl}
I_+ & := & \{(d,d') \in \mathcal{D} \times \mathcal{D} : \ dPd'\} \\
I_- & := & \{(d',d) \in \mathcal{D} \times \mathcal{D} : \ dPd'\} \\
u_{dd'} & := & \Psi(d) - \Psi(d') \in \mathbb{R}^N \qquad \forall (d,d') \in I (I := I_+ \cup I_-) \\
\Omega_P & := & \left\{\omega \in \Omega : \omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d,d' \in \mathcal{D},\ dPd'\right\} \\
\|\cdot\| & := & \gamma^\circ
\end{array}
$$

(P)

$$
\begin{array}{ll}
\min & \|\omega\|^\circ \\
\text{s.t.} & y_i(\omega^\top u_i + \beta) \geq 1 \, \forall i \in I \\
 & \omega \in \Omega
\end{array}
$$

(Q)

$$
\begin{array}{ll}
\min & \gamma(\omega) \\
\text{s.t.} & y_{dd'}(\omega^\top u_{dd'} + \beta) \geq 1 \, \forall (d,d') \in I \\
 & \omega \in \Omega
\end{array}
$$

Th: If (ω, β) : optimal to (Q), then $\beta = 0$

$$
\begin{aligned}
I_+ &:= \{(d,d') \in \mathcal{D} \times \mathcal{D} : \; dPd'\} \\
I_- &:= \{(d',d) \in \mathcal{D} \times \mathcal{D} : \; dPd'\} \\
u_{dd'} &:= \Psi(d) - \Psi(d') \in \mathbb{R}^N \qquad \forall (d,d') \in I (I := I_+ \cup I_-) \\
\Omega_P &:= \left\{\omega \in \Omega : \omega^\top \Psi(d) > \omega^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D}, \; dPd'\right\} \\
\|\cdot\| &:= \gamma^\circ
\end{aligned}
$$

(P)

$$
\begin{aligned}
\min \quad & \|\omega\|^\circ \\
\text{s.t.} \quad & y_i(\omega^\top u_i + \beta) \geq 1 \, \forall i \in I \\
& \omega \in \Omega
\end{aligned}
$$

(Q)

$$
\begin{aligned}
\min \quad & \gamma(\omega) \\
\text{s.t.} \quad & y_{dd'}(\omega^\top u_{dd'} + \beta) \geq 1 \, \forall (d,d') \in I \\
& \omega \in \Omega
\end{aligned}
$$

Th: If (ω, β) : optimal to (Q), then $\beta = 0$

Optimality conditions for (Q)

$$\begin{array}{ll} \min & \gamma(\omega) \\ \text{s.t.} & \omega^\top (\Psi(d) - \Psi(d')) \geq 1 \, \forall d, d' \in \mathcal{D}, \, dPd' \\ & \omega \in \Omega \end{array} \qquad (Q)$$

Th: Let $\omega^* \in \Omega$, feasible for (Q). Following statements are equivalent:

1. ω^* optimal for (Q)
2. $\exists \lambda = (\lambda_{dd'})_{dPd'}, \, \mu \in \mathbb{R}^J$ s.t.

$$\begin{array}{rcl} \lambda, \mu & \geq & 0 \\ \sum_{d \in \mathcal{D}} \Delta_d \Psi(d) + \sum_{j \in J} \mu_j q_j & \in & \partial\gamma(\omega^*) \\ \sum_{d' :\, dPd'} \lambda_{dd'} - \sum_{d' : d' P d} \lambda_{d'd} & = & \Delta_d \qquad \forall d \in \mathcal{D} \\ \lambda_{dd'} \left({\omega^*}^\top (\Psi(d) - \Psi(d')) - 1 \right) & = & 0 \quad \forall d, d' \in \mathcal{D}, \, dPd' \\ \mu_j \left({\omega^*}^\top q_j \right) & = & 0 \quad \forall j \in J \end{array}$$

Playing with the choice of the norm

For (Q):

Let $\omega^* \in \Omega$. Following statements are equivalent:

1. ${\omega^*}^\top \Psi(d) > {\omega^*}^\top \Psi(d') \qquad \forall d, d' \in \mathcal{D}, dPd'$
2. $\exists \gamma$, norm in $\mathbb{R}^N$, $\vartheta > 0$ s.t. $\vartheta \omega^* \in \mathcal{S}(Q)$.

Step 0: Initialize:

- Choose a norm $\gamma(\cdot)$ in $\mathbb{R}^N$.
- Take Ω_1, polyhedral cone modelling relative importance of criteria.
- Construct, by pairwise comparison among some elements of $\mathcal{D}$, P_1.
- Set $k = 1$ and go to Step 1.

Step k : Find ω^k, optimal solution to

$$
\begin{array}{ll}
\min & \gamma(\omega) \\
\text{s.t.} & \omega^\top (\Psi(d) - \Psi(d')) \geq 1 \, \forall d, d' \in \mathcal{D},\, dP_k d' \\
 & \omega \in \Omega_k,
\end{array}
$$

and find $d^k \in \mathcal{D}$, optimal to $\max_{d \in \mathcal{D}} \Psi(d)^\top \omega^k$ If (d^k, ω^k) is considered to be satisfactory then STOP with d^k as optimal solution and ω^k as vector of scores.

Else, enlarge P_k (e.g. by showing some d preferred to d^k) or reduce Ω_k. GoTo Step $k+1$.

(Toy) example: a 4-objective knapsack problem

$$\begin{array}{ll} \max & \left(r_1^\top z, \ldots, r_4^\top z\right) \\ \text{s.t.} & \omega^\top z \leq 30 \\ & z_i \in \{0,1\},\ i = 1, 2, \ldots, 10, \end{array}$$

r_1	7	6	1	9	8	2	7	7	10	4
r_2	2	1	−2	3	4	−8	2	6	9	−6
r_3	10	2	4	9	1	5	3	6	3	9
r_4	−1	4	−15	4	1	1	−6	7	−7	−1
ω	7	1	6	4	7	3	6	6	10	4

(Toy) example: a 4-objective knapsack problem

$$\begin{array}{ll} \max & (r_1^\top z, \ldots, r_4^\top z) \\ \text{s.t.} & \omega^\top z \leq 30 \\ & z_i \in \{0,1\},\ i = 1, 2, \ldots, 10, \end{array}$$

r_1	7	6	1	9	8	2	7	7	10	4
r_2	2	1	−2	3	4	−8	2	6	9	−6
r_3	10	2	4	9	1	5	3	6	3	9
r_4	−1	4	−15	4	1	1	−6	7	−7	−1
ω	7	1	6	4	7	3	6	6	10	4

(Toy) example: a 4-objective knapsack problem

$$\begin{array}{ll} \max & \left(r_1^\top z, \ldots, r_4^\top z\right) \\ \text{s.t.} & \omega^\top z \leq 30 \\ & z_i \in \{0,1\},\ i = 1,2,\ldots,10, \end{array}$$

r_1	7	6	1	9	8	2	7	7	10	4
r_2	2	1	−2	3	4	−8	2	6	9	−6
r_3	10	2	4	9	1	5	3	6	3	9
r_4	−1	4	−15	4	1	1	−6	7	−7	−1
ω	7	1	6	4	7	3	6	6	10	4

High cardinality of $\mathcal{D}$ advices against the use of any methodology which starts with a complete enumeration of $\mathcal{D}$

(Toy) example: a 4-objective knapsack problem

$$
\begin{array}{ll}
\max & (r_1^\top z, \ldots, r_4^\top z) \\
\text{s.t.} & \omega^\top z \leq 30 \\
 & z_i \in \{0,1\},\ i = 1,2,\ldots,10,
\end{array}
$$

r_1	7	6	1	9	8	2	7	7	10	4
r_2	2	1	−2	3	4	−8	2	6	9	−6
r_3	10	2	4	9	1	5	3	6	3	9
r_4	−1	4	−15	4	1	1	−6	7	−7	−1
ω	7	1	6	4	7	3	6	6	10	4

Sample some feasible solutions and define strict preferences among them, leading to a strict preference P, from which a scoring vector ω and an action d are obtained.

The single-objective knapsack problems

$$
\begin{array}{ll}
\max & r_i^\top z \\
\text{s.t.} & \omega^\top z \leq 30 \\
 & z_j \in \{0,1\} \quad \forall j
\end{array}
$$

Optimal solutions y_i

y_i											$r_i^\top y_i$
y_1	0	1	0	1	0	1	1	1	1	0	41
y_2	0	1	0	1	1	0	0	1	1	0	23
y_3	1	0	1	1	0	1	0	1	0	1	43
y_4	0	1	0	1	1	1	0	1	0	0	17

- DM: asked to sort the actions in $\{y_1, y_2, y_3, y_4\}$, thus defining $(\mathcal{D}, P_1) : y_2 P_1 y_4 P_1 y_1 P_1 y_3$
- Solving the SVM problem with linear constraints for $\Omega_1 = \mathbb{R}^N_+$ and P_1,

$$\begin{aligned} \omega^1 &= (0, 0.120879121, 0, 0.131868132) \\ d^1 &= (0, 1, 0, 1, 1, 0, 0, 1, 1, 0) \end{aligned}$$

- Such solution: not considered to be satisfactory, and P_1 is enriched, yielding P_2. In particular, the decision maker provides some feasible action, $y_5 = (1, 1, 0, 1, 0, 0, 0, 1, 1, 0)$, and $P_2 := P_1 \cup \{(y_5, d^1)\}$
- SVM Problem solved for P_2 :

$$\begin{aligned} \omega^2 &= (0, 0.199386503, 0.210122699, 0.246165644) \\ d^2 &= (1, 1, 0, 1, 1, 0, 0, 1, 0, 1). \end{aligned}$$

- d^2 : considered to be acceptable, and the process stops, yielding d^2 as optimal action, and ω^2 as vector of scores.

- DM: asked to sort the actions in $\{y_1, y_2, y_3, y_4\}$, thus defining $(\mathcal{D}, P_1) : y_2 P_1 y_4 P_1 y_1 P_1 y_3$
- Solving the SVM problem with linear constraints for $\Omega_1 = \mathbb{R}^N_+$ and P_1,

$$\begin{aligned} \omega^1 &= (0, 0.120879121, 0, 0.131868132) \\ d^1 &= (0, 1, 0, 1, 1, 0, 0, 1, 1, 0) \end{aligned}$$

- Such solution: not considered to be satisfactory, and P_1 is enriched, yielding P_2. In particular, the decision maker provides some feasible action, $y_5 = (1, 1, 0, 1, 0, 0, 0, 1, 1, 0)$, and $P_2 := P_1 \cup \{(y_5, d^1)\}$
- SVM Problem solved for P_2 :

$$\begin{aligned} \omega^2 &= (0, 0.199386503, 0.210122699, 0.246165644) \\ d^2 &= (1, 1, 0, 1, 1, 0, 0, 1, 0, 1). \end{aligned}$$

- d^2 : considered to be acceptable, and the process stops, yielding d^2 as optimal action, and ω^2 as vector of scores.

- DM: asked to sort the actions in $\{y_1, y_2, y_3, y_4\}$, thus defining $(\mathcal{D}, P_1) : y_2 P_1 y_4 P_1 y_1 P_1 y_3$
- Solving the SVM problem with linear constraints for $\Omega_1 = \mathbb{R}^N_+$ and P_1,

$$\begin{aligned} \omega^1 &= (0, 0.120879121, 0, 0.131868132) \\ d^1 &= (0, 1, 0, 1, 1, 0, 0, 1, 1, 0) \end{aligned}$$

- Such solution: not considered to be satisfactory, and P_1 is enriched, yielding P_2. In particular, the decision maker provides some feasible action, $y_5 = (1, 1, 0, 1, 0, 0, 0, 1, 1, 0)$, and $P_2 := P_1 \cup \{(y_5, d^1)\}$
- SVM Problem solved for P_2 :

$$\begin{aligned} \omega^2 &= (0, 0.199386503, 0.210122699, 0.246165644) \\ d^2 &= (1, 1, 0, 1, 1, 0, 0, 1, 0, 1). \end{aligned}$$

- d^2 : considered to be acceptable, and the process stops, yielding d^2 as optimal action, and ω^2 as vector of scores.

- DM: asked to sort the actions in $\{y_1, y_2, y_3, y_4\}$, thus defining $(\mathcal{D}, P_1) : y_2 P_1 y_4 P_1 y_1 P_1 y_3$
- Solving the SVM problem with linear constraints for $\Omega_1 = \mathbb{R}^N_+$ and P_1,

$$\begin{aligned} \omega^1 &= (0, 0.120879121, 0, 0.131868132) \\ d^1 &= (0, 1, 0, 1, 1, 0, 0, 1, 1, 0) \end{aligned}$$

- Such solution: not considered to be satisfactory, and P_1 is enriched, yielding P_2. In particular, the decision maker provides some feasible action, $y_5 = (1, 1, 0, 1, 0, 0, 0, 1, 1, 0)$, and $P_2 := P_1 \cup \{(y_5, d^1)\}$
- SVM Problem solved for P_2 :

$$\begin{aligned} \omega^2 &= (0, 0.199386503, 0.210122699, 0.246165644) \\ d^2 &= (1, 1, 0, 1, 1, 0, 0, 1, 0, 1). \end{aligned}$$

- d^2 : considered to be acceptable, and the process stops, yielding d^2 as optimal action, and ω^2 as vector of scores.

- DM: asked to sort the actions in $\{y_1, y_2, y_3, y_4\}$, thus defining $(\mathcal{D}, P_1) : y_2 P_1 y_4 P_1 y_1 P_1 y_3$
- Solving the SVM problem with linear constraints for $\Omega_1 = \mathbb{R}^N_+$ and P_1,

$$\begin{aligned} \omega^1 &= (0, 0.120879121, 0, 0.131868132) \\ d^1 &= (0, 1, 0, 1, 1, 0, 0, 1, 1, 0) \end{aligned}$$

- Such solution: not considered to be satisfactory, and P_1 is enriched, yielding P_2. In particular, the decision maker provides some feasible action, $y_5 = (1, 1, 0, 1, 0, 0, 0, 1, 1, 0)$, and $P_2 := P_1 \cup \{(y_5, d^1)\}$
- SVM Problem solved for P_2 :

$$\begin{aligned} \omega^2 &= (0, 0.199386503, 0.210122699, 0.246165644) \\ d^2 &= (1, 1, 0, 1, 1, 0, 0, 1, 0, 1). \end{aligned}$$

- d^2 : considered to be acceptable, and the process stops, yielding d^2 as optimal action, and ω^2 as vector of scores.

And that's all ...

- In conclusion ...
- Thanks 4 your attention
- Questions???

ecarrizosa@us.es

And that's all . . .

- In conclusion . . .
- Thanks 4 your attention
- Questions???

ecarrizosa@us.es