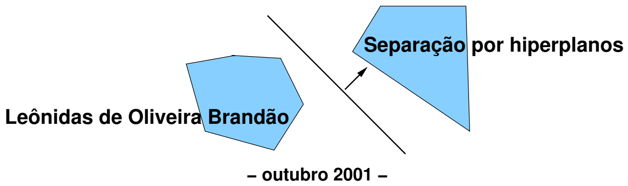


Separação de conjuntos por Hiperplanos - Perceptrons

1 Hiperplanos para separar/classificar conjuntos

1.1 Separação por Hiperplano $\mathcal{H}_{\boldsymbol{w},\boldsymbol{g}}$

- $\mathcal{A} := \{\boldsymbol{a}_1^t, \boldsymbol{a}_2^t, \ldots, \boldsymbol{a}_m^t\}$ e $\mathcal{B} := \{\boldsymbol{b}_1^t, \boldsymbol{b}_2^t, \ldots, \boldsymbol{b}_k^t\}$ (subconjuntos do $\mathbb{R}^n$)

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{bmatrix} \text{ e } \mathbf{B} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_k \end{bmatrix}$$

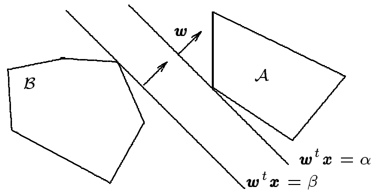
- Questão: existe $(\boldsymbol{w}, \boldsymbol{\gamma}) \in \mathbb{R}^n \times \mathbb{R}$ tal que

$$\left.\begin{array}{l} \boldsymbol{w}^t\boldsymbol{a} > \gamma, \ \forall \boldsymbol{a} \in \mathcal{A} \\ \boldsymbol{w}^t\boldsymbol{b} < \gamma, \ \forall \boldsymbol{b} \in \mathcal{B} \end{array}\right\rangle, \quad \text{ou equivalentemente,} \quad \left\langle \begin{array}{l} \mathbf{A}\boldsymbol{w} > \gamma\mathbf{1} \\ \mathbf{B}\boldsymbol{w} < \gamma\mathbf{1} \end{array}\right.$$

  - Sim $\Rightarrow$ separável (estritamente)
  - Não $\Rightarrow$ não separável

1.2 Programas lineares/não lineares

- $vo = \max_{(\boldsymbol{w},\alpha,\beta)\in\Omega} f(\boldsymbol{w}, \alpha, \beta)$
- Exemplos:
  1. $\max_{\boldsymbol{w},\alpha,\beta} \left\{\alpha - \beta : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1}, \|\boldsymbol{w}\|_\infty \leqq 1\right\}$
  2. $\min_{\boldsymbol{w},\alpha,\beta} \left\{\|\boldsymbol{w}\| - (\alpha - \beta) : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1}\right\}$

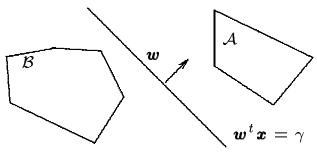


- Valor ótimo responde à questão (usando o modelo 1 acima):
  - $vo > 0 \Rightarrow$ separável $\quad (\gamma = \frac{\alpha+\beta}{2})$
  - $vo \leqq 0 \Rightarrow$ não separável

1.3 Problema de Classificação

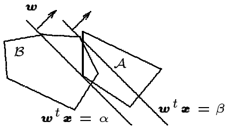
- Classificar a partir de exemplos
  - Quando separável (estrito)

$$\begin{array}{l} \boldsymbol{w}^t\boldsymbol{x} > \gamma \ \Rightarrow \ \boldsymbol{x} \in \mathcal{A} \\ \boldsymbol{w}^t\boldsymbol{x} < \gamma \ \Rightarrow \ \boldsymbol{x} \in \mathcal{B} \end{array}$$



– Ao menos, quando não separável

$\boldsymbol{w}^t\boldsymbol{x} \geqq \alpha \;\Rightarrow\; \boldsymbol{x} \in \mathcal{A}$
$\boldsymbol{w}^t\boldsymbol{x} \leqq \beta \;\Rightarrow\; \boldsymbol{x} \in \mathcal{B}$



- Encontrar separadores
  – usem poucos atributos $(\min \|\boldsymbol{x}_+\|_1,\ \min \|\boldsymbol{x}_*\|_1)$
  – generalizem bem $(\min\{\|\boldsymbol{x}_*\|_1 - \alpha - \beta\}, \max d(\mathcal{A},\mathcal{B}))$

1.4 Decidir separabilidade (estrita)

- Problema: $\mathcal{A},\mathcal{B}$ separáveis $\iff [\mathcal{A}]\bigcap[\mathcal{B}] = \emptyset$
  $\iff \big\{(\boldsymbol{\alpha},\boldsymbol{\beta})\in\mathbb{R}^n_+ \times \mathbb{R}^n_+ : \mathbf{A}'\boldsymbol{\alpha} = \mathbf{B}'\boldsymbol{\beta}, \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1\big\} = \emptyset$
- Programa não linear
  – $\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{\frac{1}{2}\|\mathbf{A}'\boldsymbol{\alpha} - \mathbf{B}'\beta\| : \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1, \boldsymbol{\alpha} \geqq \mathbf{0}, \boldsymbol{\beta} \geqq \mathbf{0}\right\}$ $(\boldsymbol{u} := \mathbf{A}'\boldsymbol{\alpha}, \boldsymbol{v} := \mathbf{B}'\boldsymbol{\beta})$
  – $vo > 0 \Rightarrow$ separável $(\boldsymbol{u} \in [\mathcal{A}], \boldsymbol{v} \in [\mathcal{B}])$
  – $vo = 0 \Rightarrow$ não separável $(\boldsymbol{u} = \boldsymbol{v} \in [\mathcal{A}]\bigcap[\mathcal{B}])$

2 Modelos não lineares

2.1 Distância entre hiperplanos separadores

- Referência: [Mangasarian 1997]
  $\mathcal{H}_{\boldsymbol{v},\boldsymbol{\theta}} := \{\boldsymbol{x} : \boldsymbol{v}'\boldsymbol{x} = \theta\}$
  – $d(\boldsymbol{p}, \mathcal{H}_{\boldsymbol{w},\boldsymbol{\beta}}) = \min_{\boldsymbol{q}\in\mathcal{H}_{\boldsymbol{v},\boldsymbol{\theta}}} \|\boldsymbol{p} - \boldsymbol{q}\|$
  – $\boldsymbol{w}_q$ tal que $\boldsymbol{w}_q \in \arg\max_{\|\boldsymbol{q}\|\leqq 1} \boldsymbol{w}'\boldsymbol{q} \Rightarrow \boldsymbol{w}'\boldsymbol{w}_q = \|\boldsymbol{w}\|'$ (norma dual)
  – $\boldsymbol{q}_p$ projeção de $\boldsymbol{p}$ em $\mathcal{H}_{\boldsymbol{w},\boldsymbol{\beta}}$
  $\boldsymbol{q}_p = \boldsymbol{p} - \lambda\boldsymbol{w}_q \Rightarrow \beta = \boldsymbol{w}'\boldsymbol{q}_p = \boldsymbol{w}'\boldsymbol{p} - \lambda\boldsymbol{w}'\boldsymbol{w}_q = \boldsymbol{w}'\boldsymbol{p} - \lambda\|\boldsymbol{w}\|'$
  logo $\lambda = \frac{\beta - \boldsymbol{w}'\boldsymbol{p}}{\|\boldsymbol{w}\|'}$
  – $\|\boldsymbol{p} - \boldsymbol{q}_p\| = \|\lambda\boldsymbol{w}_q\| = \frac{\beta - \boldsymbol{w}'\boldsymbol{p}}{\|\boldsymbol{w}\|'} \overset{\boldsymbol{p}\in\mathcal{H}_{\boldsymbol{w},\boldsymbol{\alpha}}}{=} \frac{\alpha-\beta}{\|\boldsymbol{w}\|'}$
- $\max_{\boldsymbol{w},\alpha,\beta} \left\{\frac{\alpha-\beta}{\|\boldsymbol{w}\|'} : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1}\right\}$

2.2 Dual de programa quadrático [Mangasarian]

- Dual de

$$\max\left\{-\frac{1}{2}\boldsymbol{x}'\mathbf{Q}\boldsymbol{x} + \boldsymbol{b}'\boldsymbol{y} : \mathbf{Q}\boldsymbol{x} + \mathbf{A}'\boldsymbol{y} = \boldsymbol{c}, \boldsymbol{y} \geqq \mathbf{0}\right\} \qquad (\star)$$

é

$$\min\left\{\frac{1}{2}\boldsymbol{x}'\mathbf{Q}\boldsymbol{x} - \boldsymbol{c}'\boldsymbol{x} : \mathbf{A}'\boldsymbol{y} \leqq \boldsymbol{b}\right\} \qquad (\star\star)$$

- Programa abaixo está na forma $(\star)$ [Bennett & Bredensteiner]

$$\min_{\boldsymbol{w},\alpha,\beta} \left\{\|\boldsymbol{w}\|_2^2 - (\alpha - \beta) : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, -\mathbf{B}\boldsymbol{w} \geqq -\beta\mathbf{1}\right\}$$

- Seu dual, por [Mangasarian] é

$$\max_{\boldsymbol{w},\alpha,\beta} \left\{\alpha - \beta : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, -\mathbf{B}\boldsymbol{w} \geqq -\beta\mathbf{1}, \|\boldsymbol{w}\|_2^2 = 1\right\}$$

2.3 Dual de programa quadrático [Mangasarian]

- Trocando normas $\|.\|_2$ por $\|.\|_\infty$ [Bennett & Bredensteiner]

n problemas	n problemas
$\max\ \alpha-\beta$	$\max\ \alpha-\beta$
$\mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}$	$\mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}$
$\mathbf{B}\boldsymbol{w} \geqq -\beta\mathbf{1}$	$\mathbf{B}\boldsymbol{w} \geqq -\beta\mathbf{1}$
$-\mathbf{1} \leqq \boldsymbol{w} \leqq \mathbf{1}$	$-\mathbf{1} \leqq \boldsymbol{w} \leqq \mathbf{1}$
$w_j = 1$	$w_j = -1$

3 Modelos lineares

Nesta seção mostraremos que os problemas de "maior separação" e "encontrar o ponto mais próximo" entre os convexos $[\mathcal{A}]$ e $[\mathcal{B}]$ são duais nas normas $\|.\|_1$ e $\|.\|_\infty$.

Primal	Dual
$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{\|\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}\|_\infty : \mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1, \boldsymbol{\alpha}\geqq\mathbf{0}, \boldsymbol{\beta}\geqq\mathbf{0}\right\}$	$\max_{\boldsymbol{w},\alpha,\beta}\left\{\alpha-\beta : \mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1}, \mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}, \|\boldsymbol{w}\|_1\leqq 1\right\}$ equivale a $\max_{\boldsymbol{w},\alpha,\beta}\left\{\alpha-\beta : \mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1}+\mathbf{1}, \mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}-\mathbf{1}, \|\boldsymbol{w}\|_1\leqq 1\right\}$
$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{\|\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}\|_1 : \mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1, \boldsymbol{\alpha}\geqq\mathbf{0}, \boldsymbol{\beta}\geqq\mathbf{0}\right\}$	$\max_{\boldsymbol{w},\alpha,\beta}\left\{\alpha-\beta : \mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1}, \mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}, \|\boldsymbol{w}\|_\infty\leqq 1\right\}$ equivale a $\max_{\boldsymbol{w},\alpha,\beta}\left\{\alpha-\beta : \mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1}+\mathbf{1}, \mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}-\mathbf{1}, \|\boldsymbol{w}\|_\infty\leqq 1\right\}$

3.1 Linearizando funções objetivo $\|.\|_1$ e $\|.\|_\infty$

Lema 0.1

$$(P)\ \min\left\{\|\boldsymbol{z}\|_1 : \boldsymbol{z}\in\Omega\right\}\ \Big\rangle\ \equiv\ \Big\langle (Q)\ \min\left\{\mathbf{1}'\boldsymbol{v} : -\boldsymbol{v}\leqq\boldsymbol{z}\leqq\boldsymbol{v}, \boldsymbol{z}\in\Omega\ \right\}$$

além disso,

$$\overline{\boldsymbol{z}} \text{ é solução de } (P) \iff (\overline{\boldsymbol{z}}^+, \overline{\boldsymbol{z}}) \text{ é solução de } (Q)$$

Demonstração Como

$$-\boldsymbol{v}\leqq\boldsymbol{z}\leqq\boldsymbol{v} \iff -v_i\leqq z_i\leqq v_i, \forall i \iff |z_i|\leqq v_i, \forall i,$$

é fácil ver que $\min\{\mathbf{1}'\boldsymbol{v}\} = \min\{\|\boldsymbol{z}\|_1\}$. ■

Lema 0.2

$$(P)\ \min\left\{\|\boldsymbol{z}\|_\infty : \boldsymbol{z}\in\Omega\right\}\ \Big\rangle\ \equiv\ \Big\langle (Q)\ \min\left\{\theta : -\theta\mathbf{1}\leqq\boldsymbol{z}\leqq\theta\mathbf{1}, \boldsymbol{z}\in\Omega\right\}$$

além disso,

$$\overline{\boldsymbol{z}} \text{ é solução de } (P) \iff (\|\overline{\boldsymbol{z}}\|_\infty, \overline{\boldsymbol{z}}) \text{ é solução de } (Q)$$

Demonstração A demonstração segue das seguintes equivalências

$$-\theta\mathbf{1}\leqq\boldsymbol{z}\leqq\theta\mathbf{1} \iff -\theta\leqq z_i\leqq\theta, \forall i \iff \|z_i\|\leqq\theta, \forall i \iff \|\boldsymbol{z}\|_\infty\leqq\theta,$$

$$(\theta,\boldsymbol{z})\in\mathbb{R}\times\Omega \text{ viável } (Q) \iff \|\boldsymbol{z}\|_\infty\leqq\theta,\ \text{com } (\theta,\boldsymbol{z})\in\mathbb{R}\times\Omega$$

e

$$\min\left\{\|\boldsymbol{z}\|_\infty : \boldsymbol{z}\in\Omega\right\}\ \equiv\ \min\left\{\theta : \|\boldsymbol{z}\|_\infty\leqq\theta, \boldsymbol{z}\in\Omega\right\}$$

■

Dos lemas acima, seguem os resultados

Corolário 0.1 *Os problema (P_1), (P_2) e (P_3), abaixo definidos, são equivalentes*

$(P_1)\ \min\ \|\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}\|_1$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}$	$(P_2)\ \min\ \|\boldsymbol{z}\|_1$ $\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}-\boldsymbol{z}=\mathbf{0}$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}$	$(P_3)\ \min\ \mathbf{1}'\boldsymbol{v}$ $\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}-\boldsymbol{z}=\mathbf{0}$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $-\boldsymbol{v}\leqq\boldsymbol{z}\leqq\boldsymbol{v}$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}.$

Corolário 0.2 *Os problema (P_1), (P_2) e (P_3), abaixo definidos, são equivalentes*

$(P_1)\ \min\ \|\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}\|_\infty$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}$	$(P_2)\ \min\ \|\boldsymbol{z}\|_\infty$ $\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}-\boldsymbol{z}=\mathbf{0}$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}$	$(P_3)\ \min\ \theta$ $\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}-\boldsymbol{z}=\mathbf{0}$ $\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1$ $-\theta\mathbf{1}\leqq\boldsymbol{z}\leqq\theta\mathbf{1}$ $(\boldsymbol{\alpha},\boldsymbol{\beta})\geqq\mathbf{0}.$

3.2 Programa linear para distância mínima $\|.\|_1$ entre convexos $[\mathcal{A}]$ e $[\mathcal{B}]$ separáveis

Problema não linear

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left\{\|\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}\|_1:\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1,\boldsymbol{\alpha}\geqq\mathbf{0},\boldsymbol{\beta}\geqq\mathbf{0}\right\}\tag{1}$$

Equivale ao linear

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left\{\mathbf{1}'\boldsymbol{u}+\mathbf{1}'\boldsymbol{v}:\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}-\boldsymbol{u}+\boldsymbol{v}=\mathbf{0},\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1,(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{u},\boldsymbol{v})\geqq\mathbf{0}\right\}\tag{2}$$

Encontrando o dual linear (a partir do problema canônico associado)

$$\begin{array}{l}\min \mathbf{1}'\boldsymbol{u}+\mathbf{1}'\boldsymbol{v}\\ \begin{bmatrix}\mathbf{A}' & -\mathbf{B}' & -\mathbf{I} & \mathbf{I}\\ \mathbf{1}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}'\\ \mathbf{0}' & \mathbf{1}' & \mathbf{0}' & \mathbf{0}'\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix}=\begin{bmatrix}0\\1\\1\end{bmatrix}\\ (\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{u},\boldsymbol{v})\geqq\mathbf{0}\end{array}\quad\stackrel{\text{dual}}{\longrightarrow}\quad\begin{array}{l}\max \alpha+\beta\\ \begin{bmatrix}\mathbf{A} & \mathbf{1} & \mathbf{0}\\ -\mathbf{B} & \mathbf{0} & \mathbf{1}\\ -\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{I} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{w}\\ \alpha\\ \beta\end{bmatrix}\leqq\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{1}\\ \mathbf{1}\end{bmatrix}\\ (\boldsymbol{w},\alpha,\beta)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\end{array}$$

e reescrevendo o dual, podemos simplificar o problema para a forma

$$\begin{array}{l}\max\alpha+\beta\\ \left.\begin{array}{rcl}\mathbf{A}\boldsymbol{w} & \leqq & -\alpha\mathbf{1}\\ -\mathbf{B}\boldsymbol{w} & \leqq & -\beta\mathbf{1}\end{array}\right\}\\ \left.\begin{array}{rcl}-\boldsymbol{w} & \leqq & \mathbf{1}\\ \boldsymbol{w} & \leqq & \mathbf{1}\end{array}\right\}\end{array}\quad\begin{array}{c}\stackrel{\beta\leftarrow-\beta}{\longrightarrow}\\ \\ \stackrel{\boldsymbol{w}\leftarrow-\boldsymbol{w},\,\beta\leftarrow-\beta}{\longrightarrow}\end{array}\quad\begin{array}{l}\max\alpha-\beta\\ \left\{\begin{array}{rcl}\mathbf{A}\boldsymbol{w} & \geqq & \alpha\mathbf{1}\\ \mathbf{B}\boldsymbol{w} & \leqq & \beta\mathbf{1}\end{array}\right.\\ \left\{\begin{array}{c}\\ -\mathbf{1}\leqq\ \boldsymbol{w}\ \leqq\mathbf{1}\end{array}\right.\end{array}$$

Assim, o dual de $\min_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left\{\mathbf{1}'\boldsymbol{u}+\mathbf{1}'\boldsymbol{v}:\mathbf{A}'\boldsymbol{\alpha}-\mathbf{B}'\boldsymbol{\beta}+\boldsymbol{u}-\boldsymbol{v}=\mathbf{0},\mathbf{1}'\boldsymbol{\alpha}=\mathbf{1}'\boldsymbol{\beta}=1,(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{u},\boldsymbol{v})\geqq\mathbf{0}\right\}$ fica

$$\max_{\boldsymbol{w},\alpha,\beta}\left\{\alpha-\beta:\mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1},\mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1},\|\boldsymbol{w}\|_\infty\leqq 1\right\}\tag{3}$$

- $vo>0\Rightarrow$ separável $\quad(\gamma=\frac{\alpha-\beta}{2})$
- $vo=0\Rightarrow$ não separável

Se, em (3), trocarmos as restrições $\{\mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1},\mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}\}$, por $\{\mathbf{A}\boldsymbol{w}\geqq\alpha\mathbf{1}+\mathbf{1},\mathbf{B}\boldsymbol{w}\leqq\beta\mathbf{1}-\mathbf{1}\}$, seu dual continua sendo o problema (2). Fica fácil ver isso a partir da forma

$$\begin{array}{l}\max\alpha+\beta\\ \begin{bmatrix}\mathbf{A} & \mathbf{1} & \mathbf{0}\\ -\mathbf{B} & \mathbf{0} & \mathbf{1}\\ -\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{I} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{w}\\ \alpha\\ \beta\end{bmatrix}\leqq\begin{bmatrix}-\mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\end{bmatrix}\\ (\boldsymbol{w},\alpha,\beta)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\end{array}\quad\stackrel{\text{dual}}{\longrightarrow}\quad\begin{array}{l}\min -\mathbf{1}'\boldsymbol{\alpha}+\mathbf{1}'\boldsymbol{\beta}+\mathbf{1}'\boldsymbol{u}+\mathbf{1}'\boldsymbol{v}\\ \begin{bmatrix}\mathbf{A}' & -\mathbf{B}' & -\mathbf{I} & \mathbf{I}\\ \mathbf{1}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}'\\ \mathbf{0}' & \mathbf{1}' & \mathbf{0}' & \mathbf{0}'\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix}=\begin{bmatrix}0\\1\\1\end{bmatrix}\\ (\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{u},\boldsymbol{v})\geqq\mathbf{0}\end{array}$$

e o resultado segue do fato de $(\boldsymbol{\alpha},\boldsymbol{\beta})$ serem viáveis implicar em $-\mathbf{1}'\boldsymbol{\alpha}+\mathbf{1}'\boldsymbol{\beta}=0$.

3.3 Programa linear para distância mínima $\|.\|_\infty$ entre convexos $[\mathcal{A}]$ e $[\mathcal{B}]$ separáveis

Problema não linear

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{ \|\mathbf{A}'\boldsymbol{\alpha} - \mathbf{B}'\boldsymbol{\beta}\|_\infty : \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1, \boldsymbol{\alpha} \geqq \mathbf{0}, \boldsymbol{\beta} \geqq \mathbf{0} \right\} \tag{4}$$

Equivale ao linear $\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{ \theta : \mathbf{A}'\boldsymbol{\alpha} - \mathbf{B}'\boldsymbol{\beta} - \boldsymbol{u} + \boldsymbol{v} = \mathbf{0}, \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1, -\theta\mathbf{1} \leqq \boldsymbol{u} - \boldsymbol{v} \leqq \theta\mathbf{1}, (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{v}) \geqq \mathbf{0} \right\}$

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{ \theta : \mathbf{A}'\boldsymbol{\alpha} - \mathbf{B}'\boldsymbol{\beta} - \boldsymbol{u} + \boldsymbol{v} = \mathbf{0}, \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1, \boldsymbol{u} - v + \theta\mathbf{1} - \boldsymbol{y} = \mathbf{0}, \boldsymbol{u} - v - \theta\mathbf{1} + \boldsymbol{z} = \mathbf{0}, (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{y}, \boldsymbol{z}) \geqq \mathbf{0} \right\} \tag{5}$$

Encontrando o dual linear (a partir do problema canônico associado)

$$\begin{array}{ccc}
\begin{array}{l}
\min \theta \\
\begin{bmatrix}
\mathbf{A}' & -\mathbf{B}' & -\mathbf{I} & \mathbf{I} & 0 & \mathbf{O} & \mathbf{O} \\
\mathbf{1}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & 0 & \mathbf{0}' & \mathbf{0}' \\
\mathbf{0}' & \mathbf{1}' & \mathbf{0}' & \mathbf{0}' & 0 & \mathbf{0}' & \mathbf{0}' \\
\mathbf{O} & \mathbf{O} & \mathbf{I} & -\mathbf{I} & \mathbf{1} & -\mathbf{I} & \mathbf{O} \\
\mathbf{O} & \mathbf{O} & \mathbf{I} & -\mathbf{I} & -\mathbf{1} & \mathbf{O} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{u} \\ \boldsymbol{v} \\ \theta \\ \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix}
=
\begin{bmatrix} \mathbf{0} \\ 1 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \\
(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{v}, \theta, \boldsymbol{y}, \boldsymbol{z}) \geqq \mathbf{0}
\end{array}
& \overset{\text{dual}}{\longrightarrow} &
\begin{array}{l}
\max \alpha + \beta \\
\begin{bmatrix}
\mathbf{A} & \mathbf{1} & \mathbf{0} & \mathbf{O} & \mathbf{O} \\
-\mathbf{B} & \mathbf{0} & \mathbf{1} & \mathbf{O} & \mathbf{O} \\
-\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} \\
\mathbf{I} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & -\mathbf{I} \\
\mathbf{0}' & 0 & 0 & \mathbf{1}' & -\mathbf{1}' \\
\mathbf{O} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{O} \\
\mathbf{O} & \mathbf{0} & \mathbf{0} & \mathbf{O} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix} \boldsymbol{w} \\ \alpha \\ \beta \\ \boldsymbol{r} \\ \boldsymbol{s} \end{bmatrix}
\leqq
\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ 1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \\
(\boldsymbol{w}, \alpha, \beta, \boldsymbol{r}, \boldsymbol{s}) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n
\end{array}
\end{array}$$

que pode ser simplificado como a seguir

$$\begin{array}{ccccc}
\begin{array}{rcl}
\multicolumn{3}{l}{\max\ \alpha + \beta} \\
\mathbf{A}\boldsymbol{w} + \alpha\mathbf{1} & \leqq & \mathbf{0} \\
-\mathbf{B}\boldsymbol{w} + \beta\mathbf{1} & \leqq & \mathbf{0} \\
-\boldsymbol{I}\boldsymbol{w} + \boldsymbol{r} + \boldsymbol{s} & \leqq & \mathbf{0} \\
\boldsymbol{I}\boldsymbol{w} - \boldsymbol{r} - \boldsymbol{s} & \leqq & \mathbf{0} \\
\mathbf{1}'\boldsymbol{r} - \mathbf{1}'\boldsymbol{s} & \leqq & 1 \\
-\boldsymbol{r} & \leqq & \mathbf{0} \\
\boldsymbol{s} & \leqq & \mathbf{0}
\end{array}
&
\begin{array}{c} \equiv \\ \equiv \\ \equiv \end{array}
&
\begin{array}{rcl}
\multicolumn{3}{l}{\max\ \alpha - \beta} \\
\mathbf{A}\boldsymbol{w} & \leqq & -\alpha\mathbf{1} \\
\mathbf{B}\boldsymbol{w} & \geqq & \beta\mathbf{1} \\
\boldsymbol{w} & = & \boldsymbol{r} + \boldsymbol{s} \\
\mathbf{1}'(\boldsymbol{r} - \boldsymbol{s}) & \leqq & 1 \\
\boldsymbol{r} & \geqq & \mathbf{0} \\
\boldsymbol{s} & \leqq & \mathbf{0}
\end{array}
&
\begin{array}{c}
\overset{\boldsymbol{w} \leftarrow -\boldsymbol{w},\ \beta \leftarrow -\beta}{\longrightarrow} \\
\overset{\boldsymbol{s} \leftarrow -\boldsymbol{s} \ \Rightarrow\ \boldsymbol{w}^+ = \boldsymbol{r},\ \boldsymbol{w}^- = -\boldsymbol{s})}{\longrightarrow} \\
\overset{\boldsymbol{s} \leftarrow -\boldsymbol{s} \ \Rightarrow\ \boldsymbol{w}^+ = \boldsymbol{r},\ \boldsymbol{w}^- = -\boldsymbol{s})}{\longrightarrow}
\end{array}
&
\begin{array}{rcl}
\multicolumn{3}{l}{\max\ \alpha - \beta} \\
\mathbf{A}\boldsymbol{w} & \geqq & \alpha\mathbf{1} \\
\mathbf{B}\boldsymbol{w} & \leqq & \beta\mathbf{1} \\
\boldsymbol{w} & = & \boldsymbol{w}^+ - \boldsymbol{w}^- \\
\mathbf{1}'(\boldsymbol{w}^+ - \boldsymbol{w}^-) & \leqq & 1
\end{array}
\end{array}$$

Assim, o dual de $\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{ \|\mathbf{A}'\boldsymbol{\alpha} - \mathbf{B}'\boldsymbol{\beta}\|_\infty : \mathbf{1}'\boldsymbol{\alpha} = \mathbf{1}'\boldsymbol{\beta} = 1, \boldsymbol{\alpha} \geqq \mathbf{0}, \boldsymbol{\beta} \geqq \mathbf{0} \right\}$ é

$$\max_{\boldsymbol{w},\alpha,\beta} \left\{ \alpha - \beta : \mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1}, \|\boldsymbol{w}\|_1 \leqq 1 \right\} \tag{6}$$

- $vo > 0 \Rightarrow$ separável $\quad (\gamma = \frac{\alpha - \beta}{2})$
- $vo = 0 \Rightarrow$ não separável

De modo análogo à seção anterior, se no problema (6), trocarmos as restrições $\{\mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1}\}$, por $\{\mathbf{A}\boldsymbol{w} \geqq \alpha\mathbf{1} + \mathbf{1}, \mathbf{B}\boldsymbol{w} \leqq \beta\mathbf{1} - \mathbf{1}\}$, seu dual continuará sendo o problema (5). Mais uma vez isso fica fácil a partir da primeira forma dual acima: aparecerá no lado direito, nas primeiras posições, $(\mathbf{1}; \mathbf{1})$ no lugar do $(\mathbf{0}; \mathbf{0})$, fazendo aparecer na função objetivo de seu dual $\min\{-\mathbf{1}'\boldsymbol{\alpha} + \mathbf{1}'\boldsymbol{\beta} + \theta\}$ e como $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ viáveis implica em $-\mathbf{1}'\boldsymbol{\alpha} + \mathbf{1}'\boldsymbol{\beta} = 0$ segue que seu dual é precisamente (5)

3.4 Outro programa linear

- Podemos reescrever $\mathbf{A}\boldsymbol{w} > \gamma\mathbf{1} > \mathbf{B}\boldsymbol{w}$, como

  $$\mathbf{A}\boldsymbol{w} \geqq \gamma\mathbf{1} + \mathbf{1} \ \text{ e } \ \mathbf{B}\boldsymbol{w} \leqq \gamma\mathbf{1} - \mathbf{1}$$

  $(\alpha := \min_i \boldsymbol{a}_i\boldsymbol{w}$ e $\beta := \max_i \boldsymbol{b}_i\boldsymbol{w})$

- $\min_{\boldsymbol{w},\gamma} \left\{ \frac{1}{m}(-\mathbf{A}\boldsymbol{w} + \gamma\mathbf{1} + \mathbf{1})_+ + \frac{1}{k}(\mathbf{B}\boldsymbol{w} - \gamma\mathbf{1} + \mathbf{1})_+ \right\}$

- outra forma também de [Bennett & Mangasarian]
  $\min_{\boldsymbol{w},\gamma,\boldsymbol{u},\boldsymbol{v}} \left\{ \frac{1}{m}\mathbf{1}\boldsymbol{u} + \frac{1}{k}\mathbf{1}\boldsymbol{v} : -\mathbf{A}\boldsymbol{w} + \gamma\mathbf{1} + \mathbf{1} \leqq \boldsymbol{u}, \mathbf{B}\boldsymbol{w} - \gamma\mathbf{1} + \mathbf{1} \leqq \boldsymbol{v}, (\boldsymbol{u}, \boldsymbol{v}) \geqq \mathbf{0} \right\}$