

6. Optimisation avec contraintes

méthodes primales

6.1 Approche générale

Considérons le problème général de programmation mathématique suivant:

$$\underset{x \in \tilde{R}}{\text{Min}} \quad f(x) \quad (6.1)$$

où le domaine réalisable $\tilde{R}$ est en général défini à l'aide d'un ensemble de contraintes de type $f_i(x) \leq 0, i = 1, \dots, m$ (i.e., $\tilde{R} = \{x \in R^n : f_i(x) \leq 0, i = 1, \dots, m\}$).

Considérons le problème général de programmation mathématique suivant:

$$\text{Min}_{x \in \tilde{R}} f(x). \quad (6.1)$$

Approche de résolution: **méthode itérative**

i) Itération initiale

Déterminer une solution initiale $x^0 \in \tilde{R}$.

ii) Itération générale ($k+1$)

Déterminer une direction d^k et un pas α_k dans cette direction pour s'éloigner de la solution actuelle x^k en prenant

$$x^{k+1} = x^k + \alpha_k d^k \in \tilde{R}.$$

iii) Répéter l'itération générale jusqu'à ce qu'un critère d'arrêt soit satisfait.

La direction d^k varie selon la méthode.

Par contre, les méthodes utilisent en général un "pas" optimal déterminé comme suit:

Étant donné x^k et d^k , dénotons par $\bar{\alpha}_k$ la plus grande valeur que peut prendre le pas α_k de sorte que $x^k + \bar{\alpha}_k d^k \in \tilde{R}$.

Alors le "pas" optimal α_k^* satisfait la relation

$$f\left(x^k + \alpha_k^* d^k\right) = \underset{0 \leq \alpha_k \leq \bar{\alpha}_k}{\text{Min}} \left\{ f\left(x^k + \alpha_k d^k\right) \right\} \quad (6.2)$$

6.2 Méthode des directions réalisables

Soit le problème (6.1) où $f \in C^1 / \tilde{R}$ et où $\tilde{R}$ est défini à l'aide d'un ensemble de contraintes linéaires. Ce problème prend la forme suivante:

$$\begin{aligned} (P) \quad & \text{Min } f(x) \\ & \text{Sujet à } a^{iT} x \leq b_i \quad i = 1, \dots, m. \end{aligned}$$

Étant donné une solution réalisable x^k de (P) définissons l'ensemble des indices des contraintes **actives**

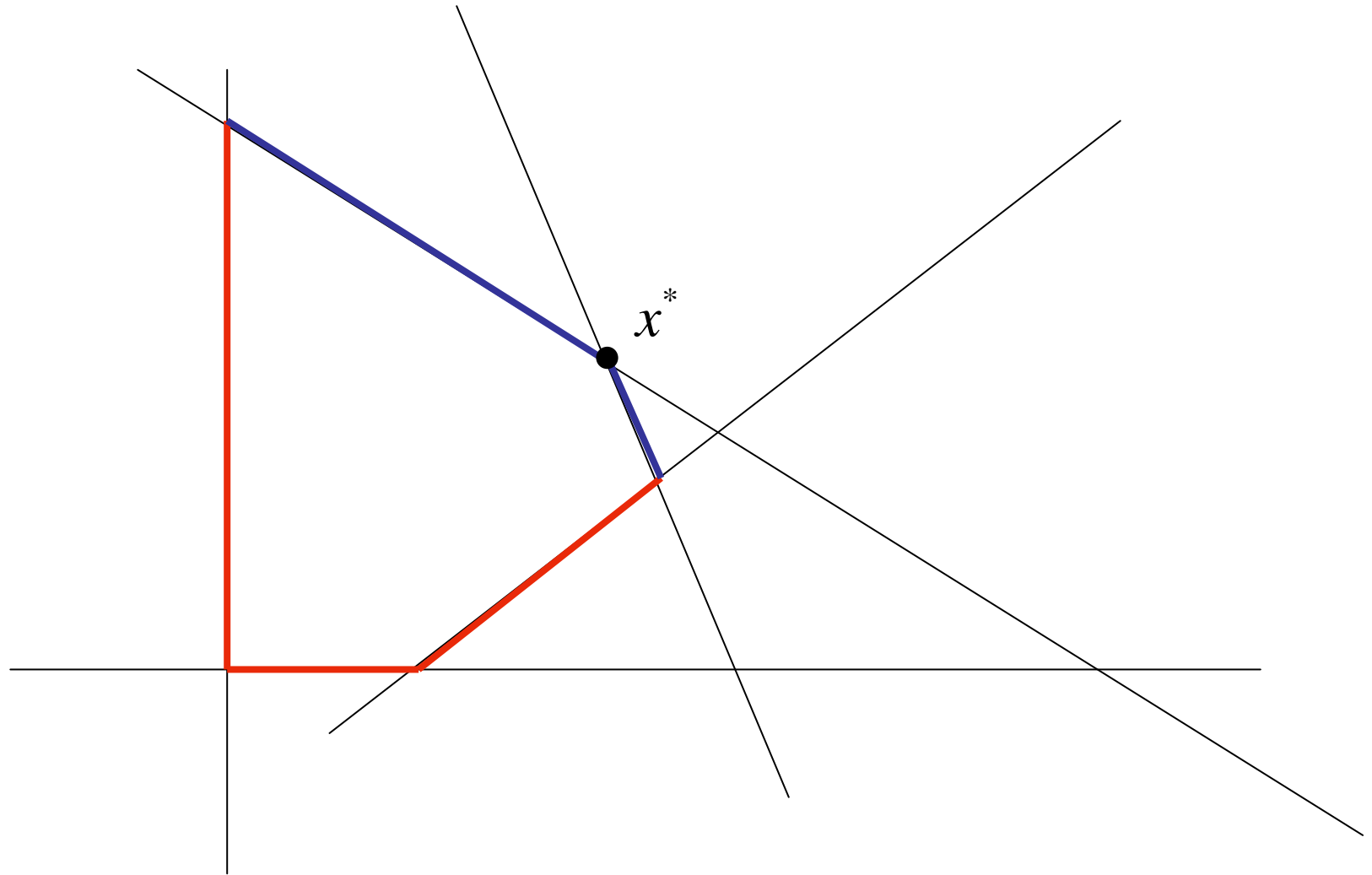
$$I_k = \left\{ i : a^{iT} x^k = b_i \right\}.$$

Puisque le pas $\alpha_k \geq 0$, alors pour que $x^k + \alpha_k d^k \in \tilde{R}$, il faut que

$$a^{iT} x^k + \alpha_k a^{iT} d^k \leq b_i \quad \forall i \in I_k$$

i.e.,

$$\alpha_k a^{iT} d^k \leq 0 \quad \forall i \in I_k.$$



$$\begin{aligned}
 (P) \quad & \text{Min } f(x) \\
 \text{Sujet à } & a^{iT} x \leq b_i \quad i = 1, \dots, m.
 \end{aligned}$$

$$I_k = \left\{ i : a^{iT} x^k = b_i \right\}.$$

Puisque le pas $\alpha_k \geq 0$, alors pour que $x^k + \alpha_k d^k \in \tilde{R}$, il faut que

$$a^{iT} x^k + \alpha_k a^{iT} d^k \leq b_i \quad \forall i \in I_k$$

i.e.,

$$\alpha_k a^{iT} d^k \leq 0 \quad \forall i \in I_k.$$

Par conséquent la direction d^k est définie à l'aide d'une solution optimale du problème

$$\begin{aligned}
 (P_k) \quad & \text{Min } \nabla f(x^k)^T d \\
 \text{Sujet à } & a^{iT} d \leq 0 \quad i \in I_k \\
 & -1 \leq d_j \leq 1 \quad j = 1, \dots, n
 \end{aligned}$$

Par conséquent la direction d^k est définie à l'aide d'une solution optimale du problème

$$\begin{array}{ll} (P_k) & \text{Min} \quad \nabla f(x^k)^T d \\ & \text{Sujet à} \quad a^{i^T} d \leq 0 \quad i \in I_k \\ & \quad \quad -1 \leq d_j \leq 1 \quad j = 1, \dots, n \end{array}$$

Les contraintes additionnelles $-1 \leq d_j \leq 1 \quad j = 1, \dots, n$ permettent de comparer les quantités $\nabla f(x^k)^T d$ sur une base équitable pour toutes les directions.

Si la valeur optimale $\nabla f(x^k)^T d^k < 0$ de (P_k) , alors la direction d^k est une direction de descente de f au point x^k et par conséquent la fonction économique f décroît dans la direction d^k dans le voisinage de x^k .

Si la valeur optimale $\nabla f(x^k)^T d^k < 0$ de (P_k) , alors la direction d^k est une direction de descente de f au point x^k et par conséquent la fonction économique f décroît dans la direction d^k dans le voisinage de x^k .

Plus grand pas $\bar{\alpha}^k$

Étant donné x^k et d^k , dénotons par $\bar{\alpha}_k$ est la plus grande valeur que peut prendre le pas α_k de sorte que $x^k + \bar{\alpha}_k d^k \in \tilde{R}$.

Ainsi, dans le cas de contraintes linéaires, $\bar{\alpha}_k$ est la plus grande valeur que peut prendre le pas α_k de sorte

$$\begin{aligned} a^{iT} (x^k + \alpha_k d^k) &\leq b_i & i = 1, \dots, m \\ a^{iT} x^k + \alpha_k a^{iT} d^k &\leq b_i & i = 1, \dots, m. \end{aligned}$$

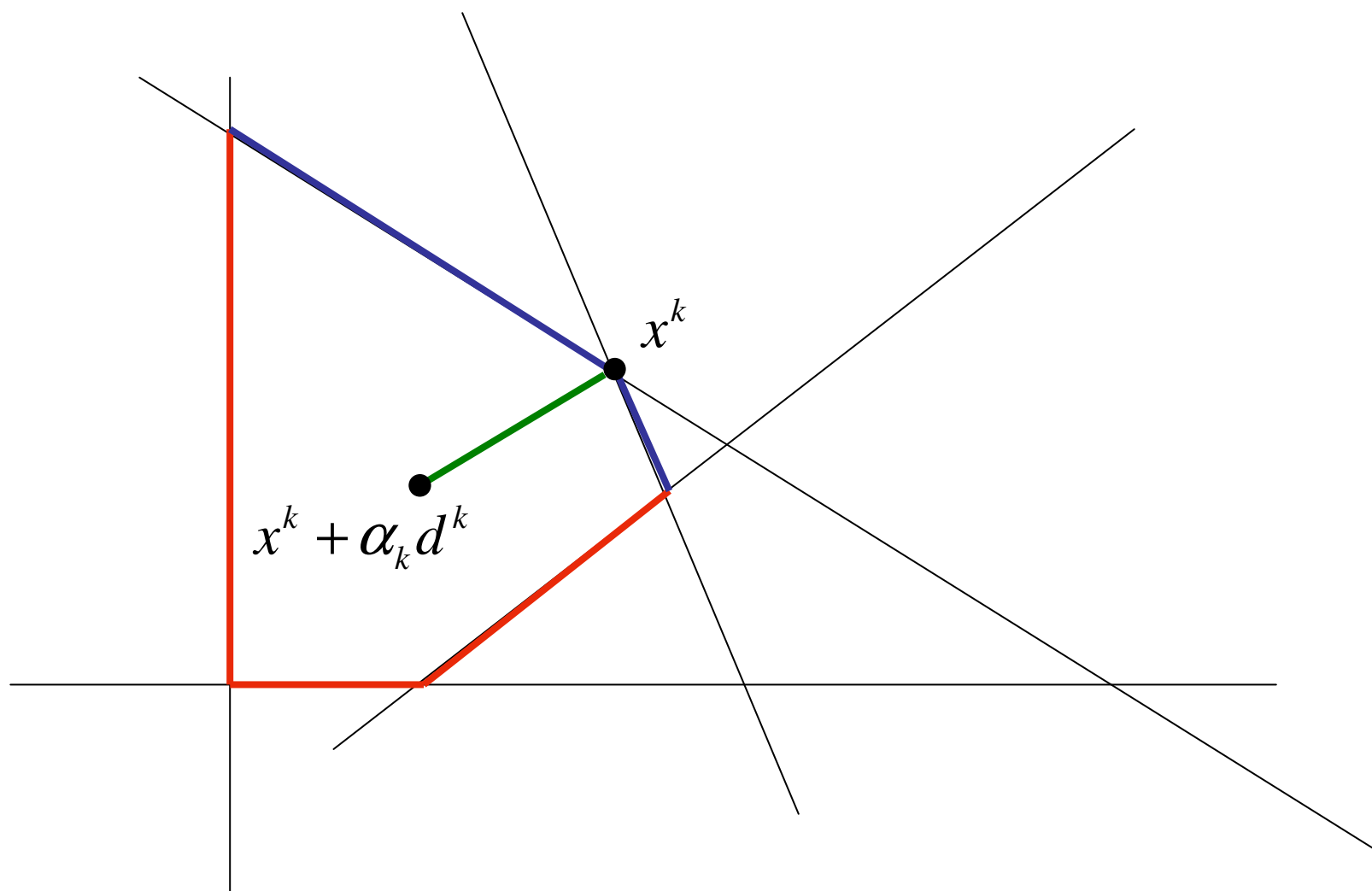
Or puisque la direction d^k a été choisi de telle sorte

$$a^{iT} (x^k + \alpha_k d^k) \leq b_i \quad i \in I_k$$

pour toute valeur de $\alpha_k \geq 0$,

alors il suffit de choisir $\bar{\alpha}_k$ comme étant la plus grande valeur que peut prendre le pas α_k de sorte

$$\begin{aligned} a^{iT} (x^k + \alpha_k d^k) &\leq b_i & i \notin I_k \\ a^{iT} x^k + \alpha_k a^{iT} d^k &\leq b_i & i \notin I_k. \end{aligned}$$



Il suffit de choisir $\bar{\alpha}_k$ comme étant la plus grande valeur que peut prendre le pas α_k de sorte

$$\begin{aligned} a^{i^T} (x^k + \alpha_k d^k) &\leq b_i & i \notin I_k \\ a^{i^T} x^k + \alpha_k a^{i^T} d^k &\leq b_i & i \notin I_k. \end{aligned}$$

Or puisque $a^{i^T} x^k < b_i$ lorsque $i \notin I_k$, il s'ensuit que

si $a^{i^T} d^k \leq 0$, alors $a^{i^T} x^k + \alpha_k a^{i^T} d^k < b_i$ pour toute valeur $\alpha_k \geq 0$;

si $a^{i^T} d^k > 0$, alors $a^{i^T} x^k + \alpha_k a^{i^T} d^k \leq b_i \Leftrightarrow \alpha_k a^{i^T} d^k \leq b_i - a^{i^T} x^k$

$$\Leftrightarrow \alpha_k \leq \frac{b_i - a^{i^T} x^k}{a^{i^T} d^k}.$$

Par conséquent

$$\bar{\alpha}_k = \text{Min}_{i \notin I_k} \left\{ \frac{b_i - a^{i^T} x^k}{a^{i^T} d^k} : a^{i^T} d^k > 0 \right\}.$$

La direction d^k varie selon la méthode.

Par contre, les méthodes utilisent en général un "pas" optimal déterminé comme suit:

Étant donné x^k et d^k , dénotons par $\bar{\alpha}_k$ la plus grande valeur que peut prendre le pas α_k de sorte que $x^k + \bar{\alpha}_k d^k \in \tilde{R}$.

Alors le "pas" optimal α_k^* satisfait la relation

$$f\left(x^k + \alpha_k^* d^k\right) = \underset{0 \leq \alpha_k \leq \bar{\alpha}_k}{\text{Min}} \left\{ f\left(x^k + \alpha_k d^k\right) \right\} \quad (6.2)$$

Le critère d'arrêt est satisfait lorsque $\nabla f(x^k)^T d^k \geq 0$ pour un certain indice k , sinon la méthode s'arrête à la limite quand $k \rightarrow \infty$.

Théorème 6.1: Supposons que f est convexe. Si la solution optimale d^k de (P_k) est telle que $\nabla f(x^k)^T d^k \geq 0$, alors x^k est une solution optimale de (P) .

Preuve. Par contradiction, supposons que x^k n'est pas une solution optimale de (P) . Alors il existe une solution réalisable $\bar{x}$ de (P) telle que $f(\bar{x}) < f(x^k)$.

Puisque f est convexe, l'inégalité du gradient s'applique et

$$f(\bar{x}) \geq f(x^k) + \nabla f(x^k)^T (\bar{x} - x^k)$$

ou

$$f(\bar{x}) - f(x^k) \geq \nabla f(x^k)^T (\bar{x} - x^k). \quad (6.3)$$

Or

$$f(\bar{x}) < f(x^k) \Rightarrow f(\bar{x}) - f(x^k) < 0$$

et substituant dans (6.3), il s'ensuit que

$$0 > f(\bar{x}) - f(x^k) \geq \nabla f(x^k)^T (\bar{x} - x^k).$$

D'autre part, si $i \in I_k$, alors $a^i{}^T x^k = b_i$ et $a^i{}^T \bar{x} \leq b_i$ et ainsi

$$a^i{}^T (\bar{x} - x^k) \leq b_i - b_i = 0.$$

$$0 > f(\bar{x}) - f(x^k) \geq \nabla f(x^k)^T (\bar{x} - x^k).$$

D'autre part, si $i \in I_k$, alors $a^{iT} x^k = b_i$ et $a^{iT} \bar{x} \leq b_i$ et ainsi

$$a^{iT} (\bar{x} - x^k) \leq b_i - b_i = 0.$$

Donc $\bar{d} = \frac{\bar{x} - x^k}{\sum_{j=1}^n |\bar{x}_j - x_j^k|}$ est une solution réalisable de (P_k) puisque

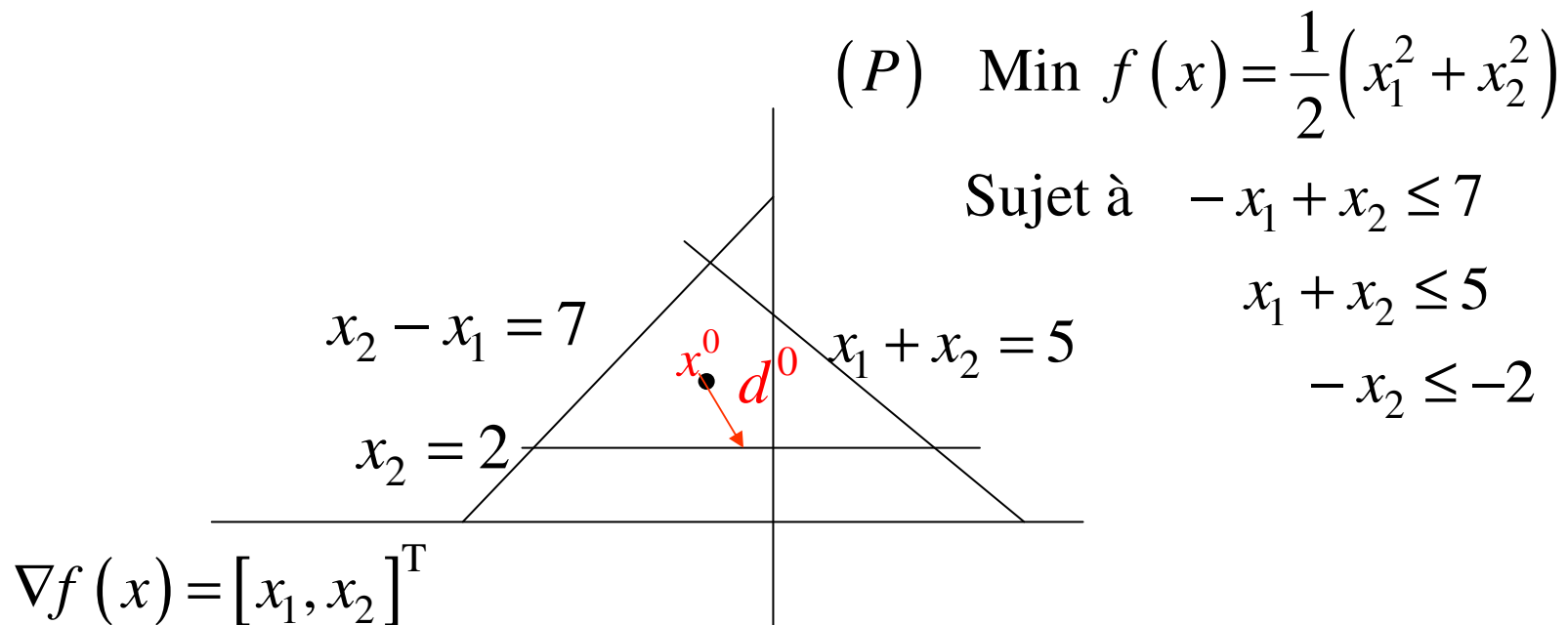
$$a^{iT} \left(\frac{\bar{x} - x^k}{\sum_{j=1}^n |\bar{x}_j - x_j^k|} \right) = \frac{1}{\sum_{j=1}^n |\bar{x}_j - x_j^k|} a^{iT} (\bar{x} - x^k) \leq 0 \quad i \in I_k.$$

$(P_k) \quad \begin{aligned} &\text{Min} \quad \nabla f(x^k)^T d \\ &\text{Sujet à} \quad a^{iT} d \leq 0 \quad i \in I_k \\ &\quad \quad \quad -1 \leq d_j \leq 1 \quad j = 1, \dots, n \end{aligned}$

Par conséquent, le fait que $\nabla f(x^k)^T \frac{\bar{x} - x^k}{\sum_{j=1}^n |\bar{x}_j - x_j^k|} < 0$ contredit l'hypothèse

que d^k est une solution optimale de (P_k) . □

Exemple numérique:



Itération 1: $x^0 = [-2, 3]^T$. Alors $I_0 = \Phi$.

La direction est solution du problème

$$\left. \begin{array}{l} (P_0) \quad \text{Min } \nabla f(x^0)^T d = -2d_1 + 3d_2 \\ \text{Sujet à } -1 \leq d_j \leq 1 \quad j = 1, 2 \end{array} \right\} \Rightarrow d^0 = [1, -1]^T$$

Exemple numérique:

$$(P) \quad \text{Min } f(x) = \frac{1}{2}(x_1^2 + x_2^2)$$

$$\text{Sujet à } -x_1 + x_2 \leq 7$$

$$x_1 + x_2 \leq 5$$

$$-x_2 \leq -2$$

$$x_2 - x_1 = 7$$

$$x_1 + x_2 = 5$$

$$x_2 = 2$$

$$x^0$$

$$d^0$$

$$\nabla f(x) = [x_1, x_2]^T$$

$$\text{Itération 1: } x^0 = [-2, 3]^T, \quad d^0 = [1, -1]^T$$

$$x^0 + \alpha_0 d^0 = [-2, 3]^T + \alpha_0 [1, -1]^T = [-2 + \alpha_0, 3 - \alpha_0]^T.$$

La plus grande valeur $\bar{\alpha}_0$:

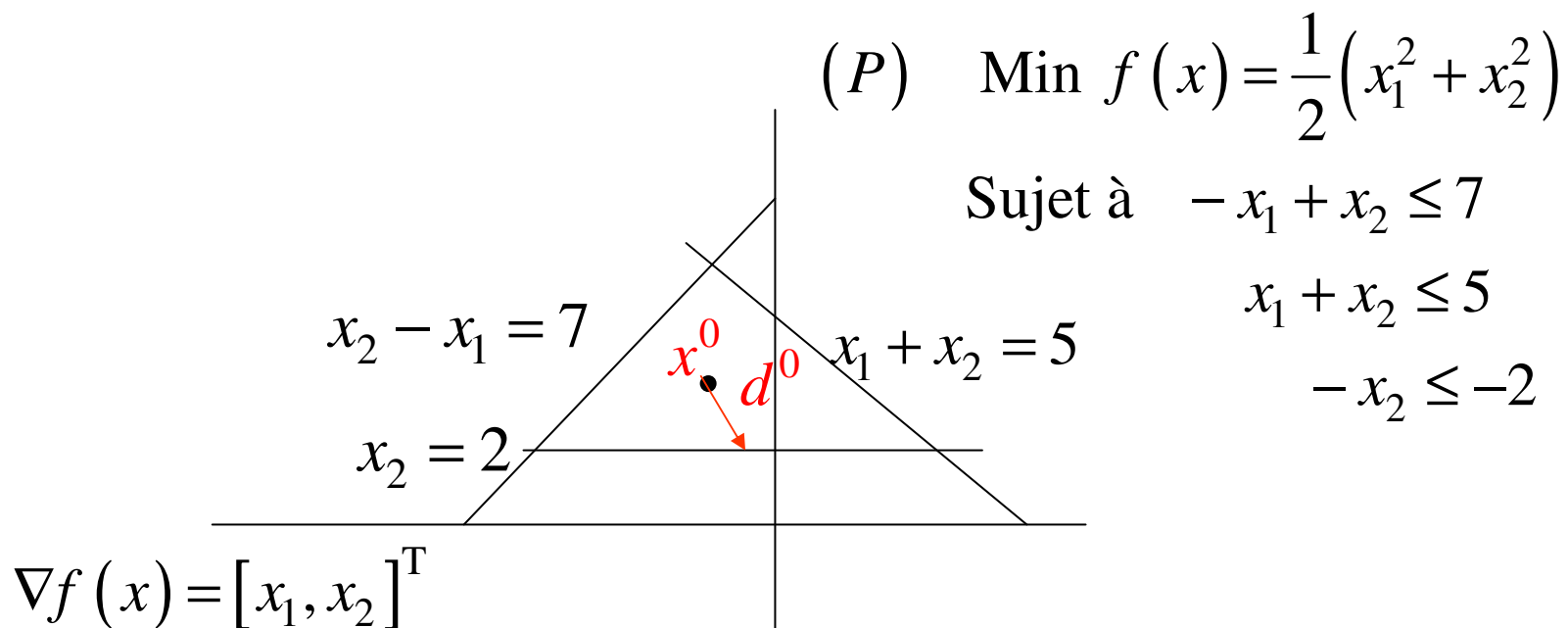
$$\left. \begin{array}{l} a^{1^T} d^0 = -1 \times 1 - 1 \times 1 = -2 < 0 \\ a^{2^T} d^0 = 1 \times 1 - 1 \times 1 = 0 \leq 0 \\ a^{3^T} d^0 = -1 \times -1 = 1 > 0 \end{array} \right\} \Rightarrow \bar{\alpha}_0 = \frac{b_3 - a^{3^T} x^0}{a^{3^T} d^0} = \frac{-2 + 3}{1} = 1$$

$$\text{si } a^{i^T} d^k > 0, \text{ alors } a^{i^T} x^k + \alpha_k a^{i^T} d^k \leq b_i \Leftrightarrow \alpha_k a^{i^T} d^k \leq b_i - a^{i^T} x^k \\ \Leftrightarrow \alpha_k \leq \frac{b_i - a^{i^T} x^k}{a^{i^T} d^k}.$$

Par conséquent

$$\bar{\alpha}_k = \text{Min}_{i \in I_k} \left\{ \frac{b_i - a^{i^T} x^k}{a^{i^T} d^k} : a^{i^T} d^k > 0 \right\}.$$

Exemple numérique:



Itération 1: $x^0 + \alpha_0 d^0 = [-2, 3]^T + \alpha_0 [1, -1]^T = [-2 + \alpha_0, 3 - \alpha_0]^T$.

La plus grande valeur $\bar{\alpha}_0$ que peut prendre α_0 pour que

$x^0 + \alpha_0 d^0 \in \tilde{R}$: lorsque la troisième contrainte devient active

$$-3 + \alpha_0 \leq -2 \Rightarrow \bar{\alpha}_0 = 1$$

Exemple numérique:

$$(P) \quad \text{Min } f(x) = \frac{1}{2}(x_1^2 + x_2^2)$$

$$\text{Sujet à } -x_1 + x_2 \leq 7$$

$$x_1 + x_2 \leq 5$$

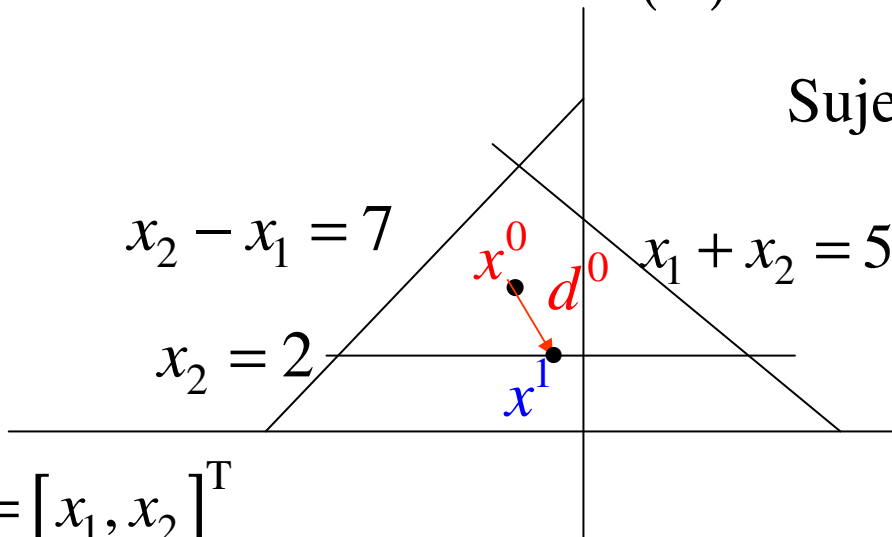
$$-x_2 \leq -2$$

$$x_2 - x_1 = 7$$

$$x_1 + x_2 = 5$$

$$x_2 = 2$$

$$\nabla f(x) = [x_1, x_2]^T$$



Itération 1: $x^0 + \alpha_0 d^0 = [-2, 3]^T + \alpha_0 [1, -1]^T = [-2 + \alpha_0, 3 - \alpha_0]^T$

$$-3 + \alpha_0 \leq -2 \Rightarrow \bar{\alpha}_0 = 1$$

$$\begin{array}{l} \text{Min}_{0 \leq \alpha_0 \leq 1} \frac{1}{2} \{13 - 10\alpha_0 + 2\alpha_0^2\} \\ \frac{d}{d\alpha_0} = -10 + 4\alpha_0 = 0 \\ \alpha_0 = \frac{5}{2} \end{array}$$

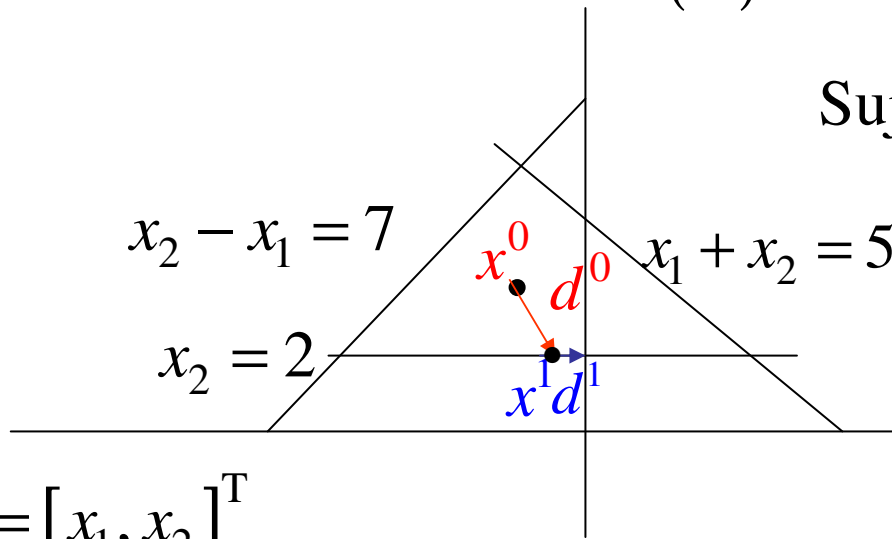
Exemple numérique:

$$(P) \quad \text{Min } f(x) = \frac{1}{2}(x_1^2 + x_2^2)$$

$$\text{Sujet à } -x_1 + x_2 \leq 7$$

$$x_1 + x_2 \leq 5$$

$$-x_2 \leq -2$$



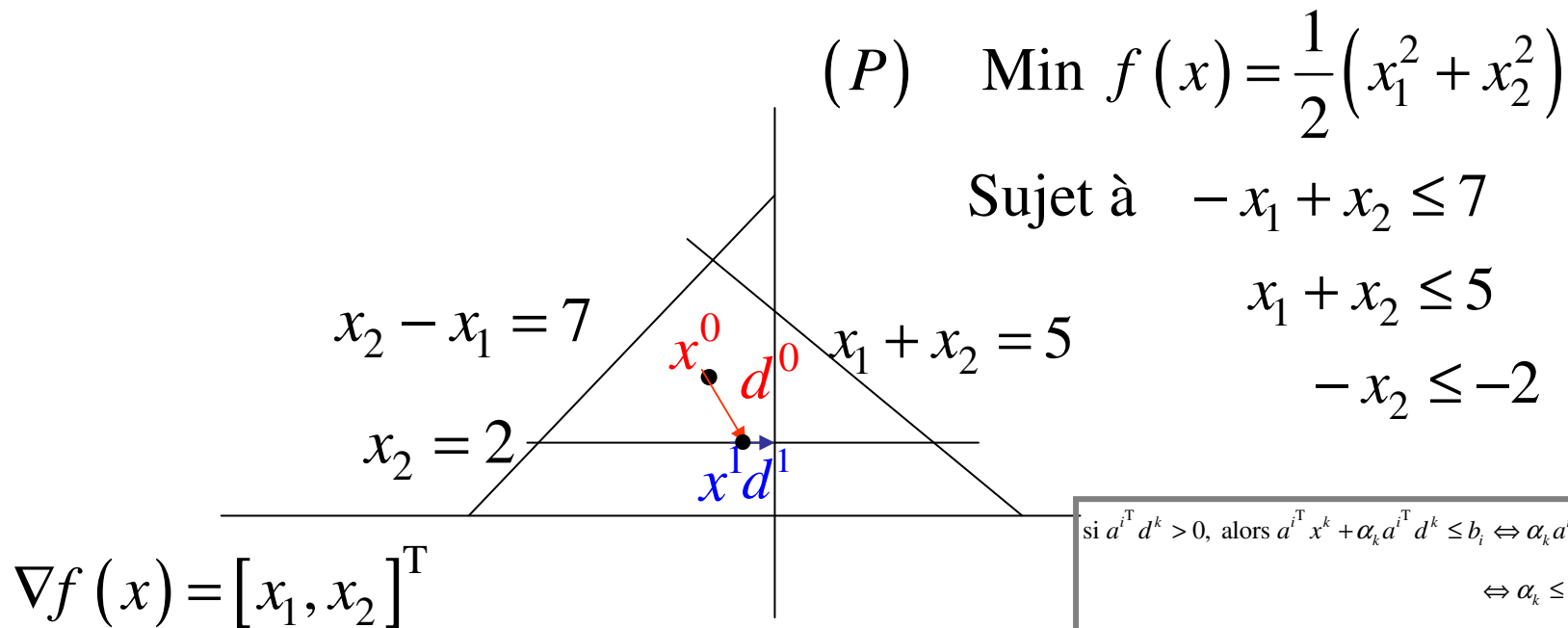
$$\nabla f(x) = [x_1, x_2]^T$$

Itération 2: $x^1 = [-1, 2]^T$. Alors $I_1 = \{3\}$.

La direction est solution du problème

$$\left. \begin{array}{l} (P_1) \quad \text{Min } \nabla f(x^1)^T d = -d_1 + 2d_2 \\ \text{Sujet à } \quad -d_2 \leq 0 \\ \quad \quad -1 \leq d_j \leq 1 \quad j = 1, 2 \end{array} \right\} \Rightarrow d^1 = [1, 0]^T$$

Exemple numérique:



$$\text{si } a^{iT} d^k > 0, \text{ alors } a^{iT} x^k + \alpha_k a^{iT} d^k \leq b_i \Leftrightarrow \alpha_k a^{iT} d^k \leq b_i - a^{iT} x^k \\ \Leftrightarrow \alpha_k \leq \frac{b_i - a^{iT} x^k}{a^{iT} d^k}.$$

Par conséquent

$$\bar{\alpha}_k = \text{Min}_{i \in I_k} \left\{ \frac{b_i - a^{iT} x^k}{a^{iT} d^k} : a^{iT} d^k > 0 \right\}.$$

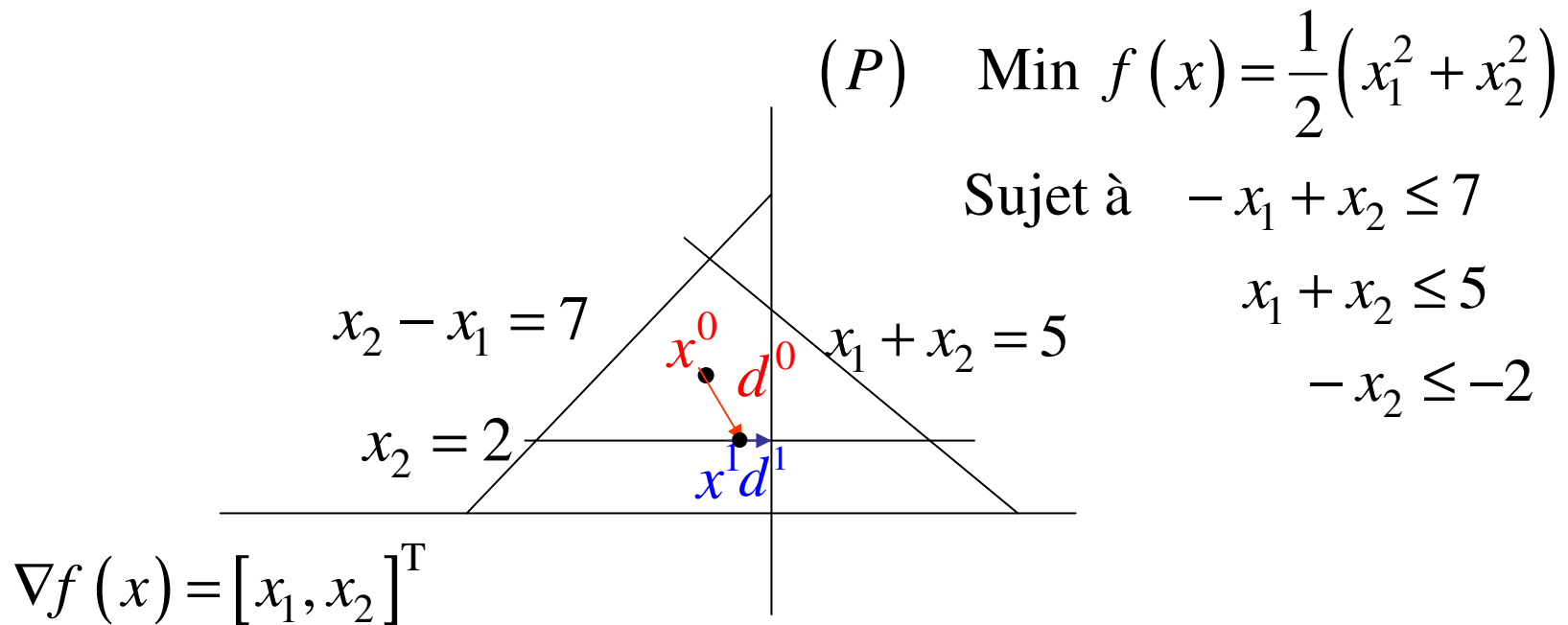
Itération 2: $x^1 = [-1, 2]^T$, $d^1 = [1, 0]^T$

$$x^1 + \alpha_1 d^1 = [-1, 2]^T + \alpha_1 [1, 0]^T = [-1 + \alpha_1, 2]^T.$$

La plus grande valeur $\bar{\alpha}_1$:

$$\left. \begin{array}{l} a^{1T} d^1 = -1 \times 1 + 0 = -1 < 0 \\ a^{2T} d^1 = 1 \times 1 + 0 = 1 > 0 \end{array} \right\} \Rightarrow \bar{\alpha}_1 = \frac{b_2 - a^{2T} x^1}{a^{2T} d^1} = \frac{5 - 1}{1} = 4$$

Exemple numérique:

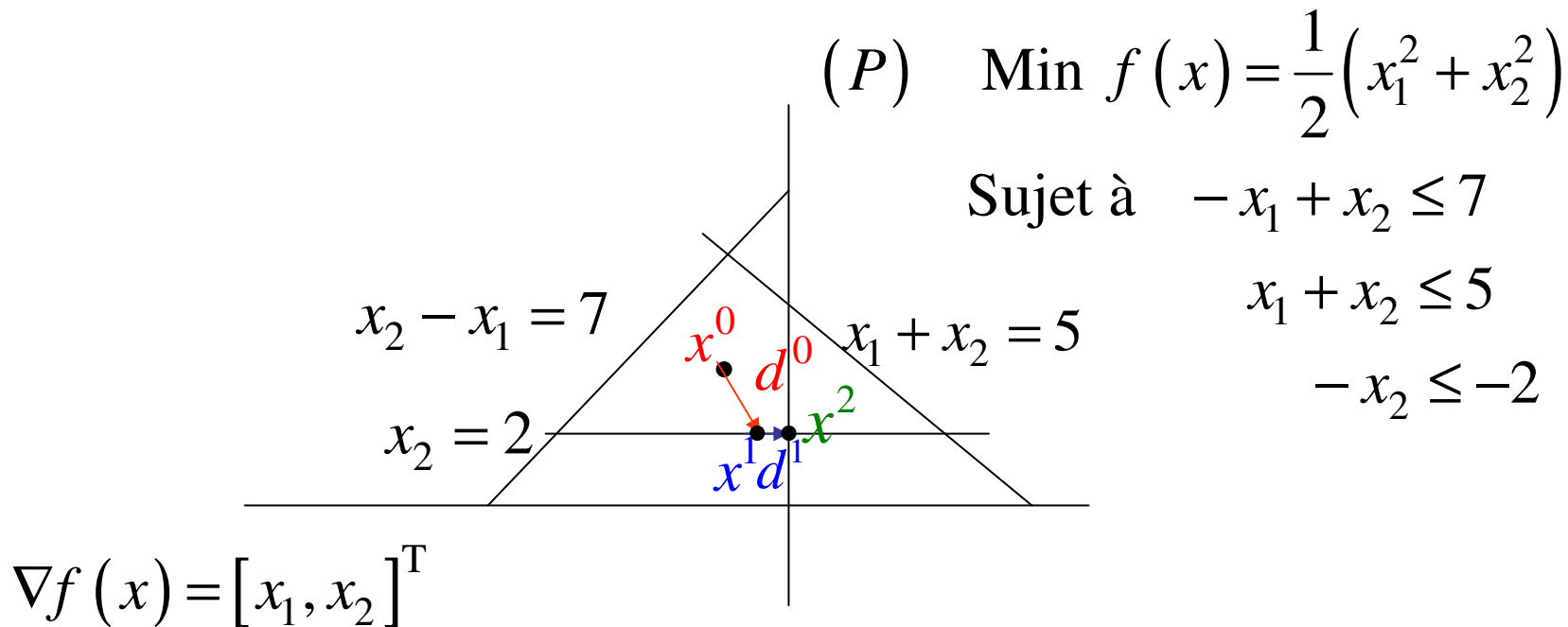


Itération 2: $x^1 + \alpha_1 d^1 = [-1, 2]^T + \alpha_1 [1, 0]^T = [-1 + \alpha_1, 2]^T$.

La plus grande valeur $\bar{\alpha}_1$ que peut prendre α_1 pour que $x^1 + \alpha_1 d^1 \in \tilde{R}$: lorsque la deuxième contrainte devient active

$$-1 + \alpha_1 + 2 \leq 5 \Rightarrow \bar{\alpha}_1 = 4$$

Exemple numérique:



Itération 2: $x^1 + \alpha_1 d^1 = [-1, 2]^T + \alpha_1 [1, 0]^T = [-1 + \alpha_1, 2]^T.$

$$-1 + \alpha_1 + 2 \leq 5 \Rightarrow \bar{\alpha}_1 = 4$$

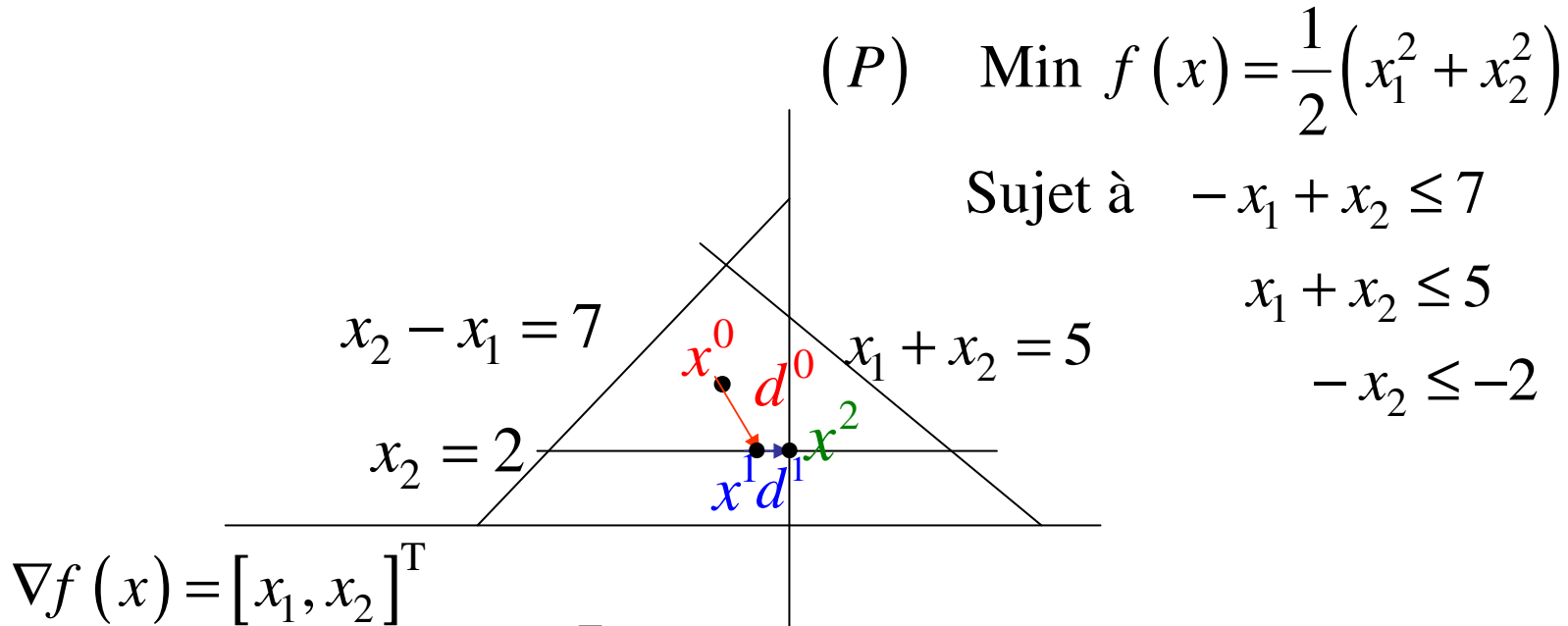
Le pas optimal α_1^* est déterminé en solutionnant le problème:

$$\text{Min}_{0 \leq \alpha_1 \leq 4} \frac{1}{2} \{ (-1 + \alpha_1)^2 + (2)^2 \} \Rightarrow \alpha_1^* = 1.$$

Donc $x^2 = [0, 2]^T.$

$$\begin{aligned} & \text{Min}_{0 \leq \alpha_1 \leq 1} \frac{1}{2} \{ 5 - 2\alpha_1 + \alpha_1^2 \} \\ & \frac{d}{d\alpha_1} = -2 + 2\alpha_1 = 0 \quad \bigg| \quad \alpha_1 = 1 \end{aligned}$$

Exemple numérique:



Itération 3: $x^2 = [0, 2]^T$. Alors $I_1 = \{3\}$.

La direction est solution du problème

$$\left. \begin{array}{l} (P_2) \quad \text{Min } \nabla f(x^2)^T d = 2d_2 \\ \text{Sujet à} \quad -d_2 \leq 0 \\ \quad \quad -1 \leq d_j \leq 1 \quad j=1,2 \end{array} \right\} \Rightarrow d^2 = [0, 0]^T$$

Donc $x^2 = [0, 2]^T$ est une solution optimale de (P).