Tutorial 7 : Reduced Gradient algorithm

Operations research, 3rd semester.

2024

Exercice 1 — ***Simple example.***

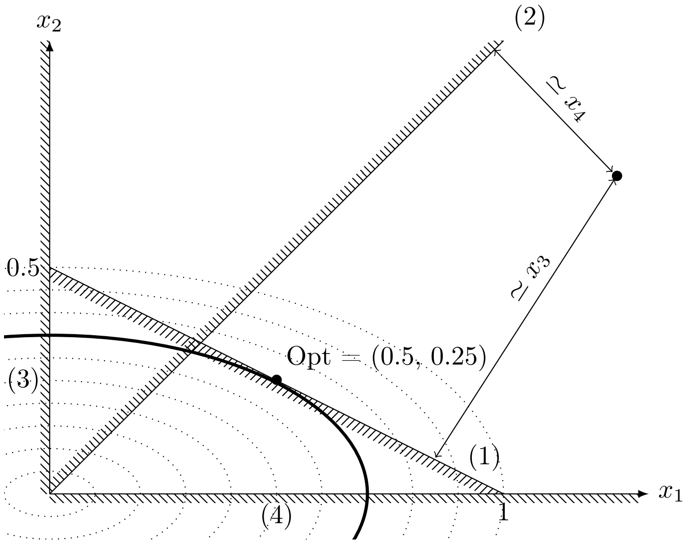
Soit (P) le problème suivant :

$$\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + 4x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + 2x_2 \geq 1 \\ -x_1 + x_2 \leq 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

1. Write the augmented form of (P) using two new variables x_3 and x_4.
2. Draw the graphical representation of the problem, include x_3 and x_4 on the drawing.
3. Apply the algorithm from $P_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ and the basis $\mathcal{B} = \{1, 3\}$ (3 iterations).
4. Check that the point satisfies the Karush-Khun-Tucker conditions.

▶ **Correction**

1. $\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + 4x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + 2x_2 - x_3 = 1 \\ -x_1 + x_2 + x_4 = 0 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$
2. The equation of an ellipse with semi-major axis a and semi-minor axis b centered at the origin is $x^2/a^2 + y^2/b^2 = 1$.
   So $x_1^2 + 4x_2^2 = R$ is the equation of an ellipse such that :
   $$\begin{cases} \frac{\sqrt{R}}{a^2} & = 1 \\ \frac{\sqrt{R}}{b^2} & = 4 \end{cases}$$
   So $a = 2b = \sqrt{R}$.
   We therefore draw ellipses around the origin that get increasingly smaller with $a = 2b$ until we exit the set of feasible solutions ; the point where we exit indicates the optimal solution.



3. We start from $P_0 = \begin{pmatrix}1\\1\\2\\0\end{pmatrix}$ with the base $B = \{1,3\}$, so $N = \{2,4\}$.

The matrix A is $\begin{pmatrix}1 & 2 & -1 & 0\\ -1 & 1 & 0 & 1\end{pmatrix}$; the rank is 2.

The gradient is $\nabla f(P_0) = \begin{pmatrix}2\\8\\0\\0\end{pmatrix}$

We compute A_B, A_N, x_B, x_N, ∇f_B and ∇f_N :

$A_B = \begin{pmatrix}1 & -1\\ -1 & 0\end{pmatrix}, \quad A_N = \begin{pmatrix}2 & 0\\ 1 & 1\end{pmatrix}$

$x_B = \begin{pmatrix}1\\2\end{pmatrix}, \quad x_N = \begin{pmatrix}1\\0\end{pmatrix}$

$\nabla f_B = \begin{pmatrix}2\\0\end{pmatrix}, \quad \nabla f_N = \begin{pmatrix}8\\0\end{pmatrix}$

The matrix A_B is invertible and $A_B^{-1} = \begin{pmatrix}0 & -1\\ -1 & -1\end{pmatrix}$.

We can compute the reduced gradient :

$$\begin{aligned}{}^t\nabla \bar{f}(x_N) &= -{}^t\nabla f_B \cdot A_B^-1 \cdot A_N + {}^t\nabla f_N\\ &= \begin{pmatrix}2 & 0\end{pmatrix} \cdot \begin{pmatrix}1 & 1\\ 3 & 1\end{pmatrix} + \begin{pmatrix}8 & 0\end{pmatrix}\\ &= \begin{pmatrix}10 & 2\end{pmatrix}\end{aligned}$$

Second method : explicitely compute $\bar{f}(x_N)$.

$$\begin{aligned}f(x) &= x_1^2 + 4x_2^2\\ \bar{f}(x_N) &= (x_2 + x_4)^2 + 4x_2^2\\ &= 5x_2^2 + x_4^2 + 2x_2x_4\end{aligned}$$

So $\nabla \bar{f}(x_N) = \begin{pmatrix}10x_2 + 2x_4\\ 2x_4 + 2x_2\end{pmatrix} = \begin{pmatrix}10\\2\end{pmatrix}$

We compute now the direction $d_N = \begin{pmatrix} -10 \\ 0 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_2} > 0 \text{ et } x_2 \neq 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} > 0 \text{ et } x_4 = 0 \end{matrix}$

We deduce $d_B = -A_B^{-1} A_N d_N = \begin{pmatrix} -10 \\ -30 \end{pmatrix} \begin{matrix} (x_1) \\ (x_3) \end{matrix}$

The direction is not null, we search then $\alpha_1 =$ the maximum $\alpha \geq 0$ such that $P_0 + \alpha \cdot d \geq 0$.

We have then $\begin{cases} 1 - 10\alpha & \geq 0 \\ 1 - 10\alpha & \geq 0 \\ 2 - 30\alpha & \geq 0 \\ 0 - 0\alpha & \geq 0 \end{cases}$

So $\alpha_1 = \frac{1}{15}$.

We are now looking for the α between 0 and α_1 that minimizes $f(P_0 + \alpha d)$.

$$\begin{aligned} g(\alpha) = f(P_0 + \alpha d) &= (1 - 10\alpha)^2 + 4(1 - 10\alpha)^2 \\ &= 5(1 - 10\alpha)^2 \\ g'(\alpha) &= -100(1 - 10\alpha) \\ g'(\alpha) &\leq 0 \\ \Leftrightarrow \alpha &\leq \frac{1}{10} \end{aligned}$$

The function g is therefore decreasing between 0 and $\alpha_1 \leq \frac{1}{10}$, we set $\alpha_2 = \alpha_1 = \frac{1}{15}$.

We move to $P_1 = P_0 + \frac{1}{15} d = \begin{pmatrix} 1/3 \\ 1/3 \\ 0 \\ 0 \end{pmatrix}$

We have x_3 that is zero. Therefore, we need to change the basis. We try to replace x_3 with x_2 in the basis (since $x_2 \geq x_4$).

Let $B' = \{1, 2\}$, We have $A_{B'} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$ of determinant 3, hence invertible. We can therefore perform this change of basis.

We now have $B = \{1, 2\}$ and $N = 3, 4$.

The gradient is $\nabla f(P_0) = \begin{pmatrix} 2/3 \\ 8/3 \\ 0 \\ 0 \end{pmatrix}$

We compute A_B, A_N, x_B, x_N, ∇f_B and ∇f_N :

$A_B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \quad A_N = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

$x_B = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}, \quad x_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\nabla f_B = \begin{pmatrix} 2/3 \\ 8/3 \end{pmatrix}, \quad \nabla f_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

A_B is invertible and $A_B^{-1} = \begin{pmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{pmatrix}$.

One can thus calculate the reduced gradient :

$$\begin{aligned} {}^t\nabla \bar{f}(x_N) &= -{}^t\nabla f_B \cdot A_B^- 1 \cdot A_N + {}^t\nabla f_N \\ &= \begin{pmatrix} 2/3 & 8/3 \end{pmatrix} \cdot \begin{pmatrix} -1/3 & -1/3 \\ -2/3 & 1/3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 10/9 & -4/9 \end{pmatrix} \end{aligned}$$

We compute the direction $d_N = \begin{pmatrix} 0 \\ 4/9 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_3} > 0 \text{ et } x_3 = 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} < 0 \text{ et } x_4 \neq 0 \end{matrix}$

We deduce $d_B = -A_B^{-1} A_N d_N = \begin{pmatrix} 8/27 \\ -4/27 \end{pmatrix} \begin{matrix} (x_1) \\ (x_2) \end{matrix}$

The direction is non-zero, so we are looking for $\alpha_1 =$ the maximum $\alpha \geq 0$ such that $P_1 + \alpha \cdot d \geq 0$.

We then have $\begin{cases} 1/3 + 8/27\alpha & \geq 0 \\ 1/3 - 4/27\alpha & \geq 0 \\ 0 + 0\alpha & \geq 0 \\ 0 + 4/9\alpha & \geq 0 \end{cases}$

So $\alpha_1 = \frac{9}{4}$.

We are now looking for the α between 0 and α_1 that minimizes $f(P_0 + \alpha d)$.

$$\begin{aligned} g(\alpha) = f(P_0 + \alpha d) &= (1/3 + 8/27\alpha)^2 + 4(1/3 - 4/27\alpha)^2 \\ g'(\alpha) &= 2 * 8/27 * (1/3 + 8/27\alpha) - 8 * 4/27 * (1/3 - 4/27\alpha) \\ g'(\alpha) &= -1/3 + 16/27 \cdot \alpha \\ g'(\alpha) &\leq 0 \\ \Leftrightarrow \alpha &\leq \frac{9}{16} \end{aligned}$$

The function g is therefore decreasing between 0 and $\frac{9}{16} \leq \alpha_1$, we set $\alpha_2 = \frac{9}{16}$.

We move to $P_2 = P_1 + \frac{9}{16} d = \begin{pmatrix} 1/2 \\ 1/4 \\ 0 \\ 1/4 \end{pmatrix}$

We do not have $x_1 = 0$ or $x_2 = 0$, we keep the same basis.

We compute x_B, x_N, ∇f_B and ∇f_N :

$x_B = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}, \quad x_N = \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}$

$\nabla f_B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \nabla f_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

We compute the reduced gradient :

$$\begin{aligned} {}^t\nabla \bar{f}(x_N) &= -\,{}^t\nabla f_B \cdot A_B^{-}1 \cdot A_N + {}^t\nabla f_N \\ &= \begin{pmatrix} 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1/3 & -1/3 \\ -2/3 & 1/3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \end{aligned}$$

We compute the direction $d_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_3} > 0 \text{ et } x_3 = 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} = 0 \text{ et } x_4 \neq 0 \end{matrix}$

The direction d_N is zero, so we stop and the (KKT) conditions are satisfied.

4. We look for $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0, \mu_1, \mu_2 \in \mathbb{R}$ such that

$$\begin{cases} \nabla f(P_2) + \sum_{i=1}^{4} \lambda_i \nabla g_i(P_2) + \sum_{i=1}^{2} \mu_j \nabla h_j(P_2) & = 0 \\ \lambda_i g_i(P_2) & = 0 \ \forall i \leq 4 \end{cases}$$

$$\begin{cases} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & = 0 \\ \lambda_1 1/2 & = 0 \\ \lambda_2 1/4 & = 0 \\ \lambda_3 0 & = 0 \\ \lambda_4 11/4 & = 0 \end{cases}$$

$$\begin{cases} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & = 0 \\ \lambda_1 & = 0 \\ \lambda_2 & = 0 \\ \lambda_4 & = 0 \end{cases}$$

We set $\mu_2 = 0$, $\mu_1 = -1$ and $\lambda_3 = 1$ and the system is then satisfied.

Exercice 2 — ***Reduced gradient and standard form***

Let (P) be the following program.

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.c.} \quad \begin{cases} \sum_{j=1}^{n} a_{ij} x_j \leq b_i \ \forall i \in [\![1; m]\!] \\ x \geq 0 \end{cases}$$

We assume that the functions g_i are linear and that there exists a feasible solution $z > 0$.

1. Write the augmented form of (P) by adding m variables. Let $y_1, y_2, \dots, y_{n+m}$ be the variables of that new program and f' be the new objective function.
2. Show that if we set $\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ to z then, for every $i \in [\![1; m]\!]$, $y_{n+i} \geq 0$. Let y_0 be that solution.
3. Show that it is possible to start the reduced gradient algorithm from y_0 with the basis $B = \{n + i, i \in [\![1; m]\!]\}$.
4. Compute the direction d obtained at the first iteration and show that $i \leq n$, $d_i = -(\nabla f(z))_i$.

► **Correction**

1. The augmented form

$$\min_{y \in \mathbb{R}^{n+m}} f'(y) \quad \text{s.c.} \quad \begin{cases} \sum_{j=1}^{n} a_{ij} y_j + y_{n+i} = b_i \ \forall i \in [\![1; m]\!] \\ y \geq 0 \end{cases}$$

2. Let $i \leq m$, then $\sum_{j=1}^{n} a_{ij} z_j \leq b_i$ or $\sum_{j=1}^{n} a_{ij} z_j + y_{n+i} = b_i$ then $y_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} z_j \geq 0$.
3. y_0 is a feasible solution and the matrix A_B associated with this basis is the identity, so it is invertible. We can therefore start the algorithm from this point with this basis.

4. $A_B = I_m$, A_N is the matrix with coefficients a_{ij}.
   $x_N = x_0$
   $\nabla f'_N : (\nabla f'(y_0))_i = (\nabla f(x_0))_i$ if $i \in N$
   $\nabla f'_B : (\nabla f'(y_0))_i = 0$ if $i \in B$
   so ${}^t\nabla \bar{f}'(x_N) = -A_N \nabla f'_B + \nabla f'_N = \nabla f'_N$.
   For all $i \in \mathbb{N}$, in other words, $i \leq n$, we have $y_i > 0$. Therefore, $d_i = -\nabla \bar{f}'(x_N)_i = -(\nabla f_N)_i = -(\nabla f(x_0))_i$.

Exercice 3 — *Linear objective*

Let (P) be the following program :

$$\min_{x\in\mathbb{R}^n} c\cdot x \quad \text{s.c.} \quad \begin{cases} Ax = b \\ x \geq 0 \end{cases}$$

where A is a $m\times n$ matrix with $m \leq n$ and where b is a vector of size m.

We can solve such a program with the simplex algorithm, which is similar to the reduced gradient algorithm. We recall that each non-basic variable is nul.

1. Let x be a basic feasible solution : a feasible solution and a basis B such that for all $i \in N$, $x_N = 0$. Compute the reduced cost gradient and, this way, find the formula of the reduced costs of the simplex algorithm.
2. We assume that $Ax = b; x \geq 0$ is a bounded space. Show that, during an iteration of the reduced gradient algorithm, if $d \neq \vec{0}$, there is necessarily a variable of x that is nul at the end of the iteration.
3. Is there necessarily a change of basis in that case ?
4. We recall that the simplex algorithm moves, at each iteration, by interverting two variables from the basis and the non-basis. Show that, even if we start at the same point with the same basis, there exists cases where the reduced gradient algorithm and the simplex algorithm choose different directions.

► Correction

1. We have
   $x_N = 0$
   $\nabla f_N = \left(c_i\right)_{i\in N}$ and $\nabla f_B = \left(c_i\right)_{i\in B}$
   so ${}^t\nabla \bar{f}'(x_N) = -{}^t\nabla f_B \cdot A_B^{-1}\cdot A_N + {}^t\nabla f_N$.
   We write $\bar{A} = A_B^{-1}\cdot A_N$ a matrix of size $|B|\times|N|$
   so $\nabla \bar{f}'(x_N)_i = -({}^t\nabla f_B \cdot \bar{A})_i + (\nabla f_N)_i$
   so $\nabla \bar{f}'(x_N)_i = -(\sum_{j\in B} \bar{a}_{ji}c_j) + c_i$
   We get then the reduced costs
2. If $d \neq \vec{0}$, then we follow the direction $x + \alpha d$ with :
   — $\alpha_1 = \max\{\alpha | x + \alpha d \geq 0\}$
   — $\alpha_2 = \arg\min\{c\cdot(x+\alpha d) | 0 \leq \alpha \leq \alpha_1\}$
   However $c\cdot(x+\alpha d) = c\cdot x + \alpha c\cdot d$. Let us show that $c\cdot d < 0$.
   We then have the function that is decreasing with respect to α, so we necessarily have $\alpha_1 = \alpha_2$.
   We can show that there exists $d_i < 0$ for at least one i. Otherwise, for all $\alpha \geq 0$, $x + \alpha_1 d > 0$; however, we know that $A(x+\alpha d) = b$ (result from the course). Thus, we would have $\alpha_1 = +\infty$, which would contradict the assumption that $Ax = b; x \geq 0$ is bounded.
   Let us set $i = \arg\min\{-x_j/d_j | j \leq n, d_j < 0\}$; then $\alpha_1 = -x_i/d_i$ and $(x+\alpha_1 d)_j = 0$. Therefore, the i-th variable x_i becomes zero at the end of the iteration.
3. There is not necessarily a change of basis.
   For example, with the following program :
   $$\min_{x\in\mathbb{R}^n} x_1 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 + x_3 = 1 \\ x \geq 0 \end{cases}$$
   We start from $(0.25, 0.25, 0.5)$. We use the basis $B = \{3\}$. We compute the direction $d = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ (look at exercice 2 to do this calculation quickly)..
   We follow the direction and move to $\begin{pmatrix} 0 \\ 0.25 \\ 0.75 \end{pmatrix}$. There is no change of basis because x_1 is not in the basis, so its cancellation does not change anything.

4. We use the following program as an example

$$\min_{x\in\mathbb{R}^n} -x_1 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 + x_3 = 4 \\ -x_1 + x_2 + x_4 = 1 \\ x_1 + x_2 - x_5 = 2 \\ x \geq 0 \end{cases}$$

In this program, every point visited by the simplex has at least 2 zero coordinates.
We start from the point $(0.5, 1.5, 2, 0, 0)$. We take the basis $B = \{1, 2, 3\}$. Indeed, we have $x_N = 0$. We calculate the direction and find reduced costs of -0.5 for x_4 and x_5. $d = \begin{pmatrix} 0.5 \\ 0 \\ -0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$.

We move in the given direction, we have $\alpha_1 = \alpha_2 = 4$, and we arrive at the point. $\begin{pmatrix} 2.5 \\ 1.5 \\ 0 \\ 2 \\ 2 \end{pmatrix}$.

We are no longer at a point with 2 zero coordinates, so it cannot be the direction chosen by the simplex algorithm.