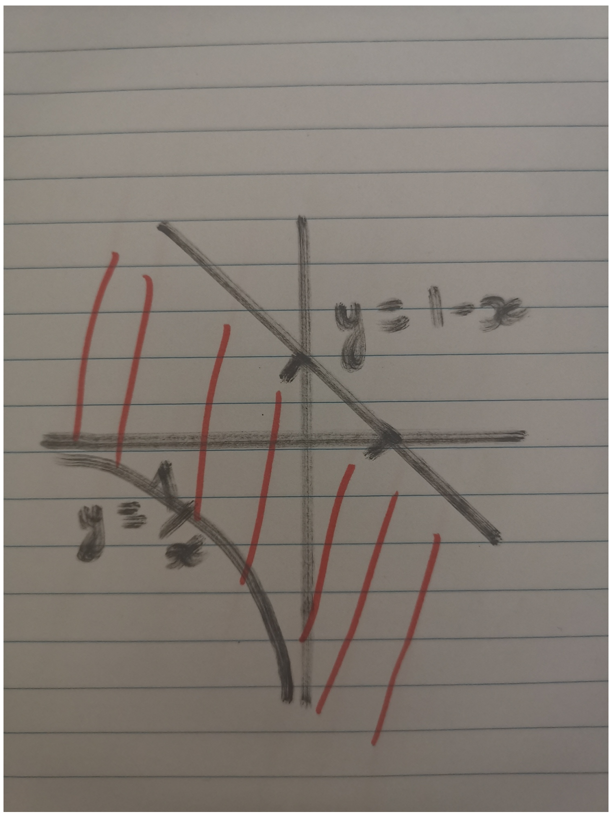
# Examen session 1: optimisation numérique (correction)

## Exercice 1

1. (1 pt)




2. (2pts)
   En $(0,0)$, on a $\nabla J(0,0) = (0,0)$ et

$$HJ(0,0) = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}$$

   Cette matrice a deux valeurs propres de signe opposé. Elle n'est donc pas positive, ce qui contredit la condition nécessaire d'optimalité d'ordre 2.

3. (2 pts)
   J est coercive sur $D$ car si $|x| \geq 1$, on a $J(x) \geq x^2+y^2+2xy$ et si $|x| \leq 1$, J(x,y)$\geq y^2-|y|$. Dans les deux cas $J(x,y)$ tend vers $+\infty$ quand $||(x,y)||$tend vers $+\infty$.

   De plus, $D$ est fermé (inégalités larges), ce qui implique bien que $J$ possède au moins un minimum global sur $D$.

4. (3 pts)

Les relations KKT s'écrivent :

$$\begin{cases} 4x^3 + 2y + \mu_1 y + \mu_2 = 0 \\ 2y + 2x + \mu_1 x + \mu_2 = 0 \\ \mu_1 \geq 0 \\ \mu_2 \geq 0 \\ x + y \leq 1 \\ xy \leq 1 \\ \mu_1(xy - 1) = 0 \\ \mu_2(x + y - 1) = 0 \end{cases}$$

On traite les 4 cas suivants :

— $\mu_1 = 0$ et $\mu_2 = 0$. Dans ce cas, on trouve que $(x, y) = (0, 0)$, $(x, y) = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ ou $(x, y) = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

— $\mu_1 = 0$ et $\mu_2 > 0$. Dans ce cas, on trouve $\mu_2 = -2(x + y) = -2$ ce qui est absurde.

— $\mu_1 > 0$ et $\mu_2 = 0$. Dans ce cas, on trouve $\mu_1 = -2 - 2y^2 < 0$ ce qui est absurde.

— $\mu_1 > 0$ et $\mu_2 > 0$. Ce cas est impossible (pas de solution au système $x + y = 1$ et $xy = 1$.

Au final, comme le point $(0, 0)$ n'est pas un minimum local, ce sont les deux points $(x, y) = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ et $(x, y) = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, qui ont la même valeur par $J$, qui sont les minima locaux (et globaux) de $J$ sur $D$.

**Exercice 2**

1. (2 pts)
   La fonction $f(\rho) : \rho \mapsto J(x_k - \rho d_k)$ est une fonction quadratique et coercive à une variable (car restriction de $J$) . Son minimum est donc son unique point critique. Comme

$$f(\rho) = \frac{1}{2} < Ad_k, d_k > \rho^2 - < x_k, Ad_k > \rho + \rho < b, d_k > + Cste$$

   on a

$$< Ad_k, d_k > \rho_k - < Ax_x, d_k > + < b, d_k > = 0$$

   soit, comme $d_k = Ax_k - b$,

$$\rho_k = \frac{||d_k||^2}{< Ad_k, d_k >}$$

2. (2 pts)
   Cela peut se faire avec l'expression de $J$ et de $\rho_k$ ou de manière générale en écrivant autrement $f'(\rho_k)$ :

$$0 = f'(\rho_k) = < d_k, \nabla J(x_k - \rho_k d_k) > = < d_k, d_{k+1} >$$

3. (2 pts)
   On déduit de ces deux inégalités :

$$||x_k - x^*||^2 \leq \frac{\lambda_n^2}{2\lambda_1}(J(x_k) - J(x_{k+1}))$$

   En sommant ces inégalités de 1 à $N$, il vient

$$\sum_{k=1}^{N} ||x_k - x^*||^2 \leq \frac{\lambda_n^2}{2\lambda_1}(J(x_0) - J(x_{N+1}))$$

   Comme $J$ est minorée, on en déduite que la série de terme général positif $u_k = ||x_k - x^*||^2$ est convergente et donc que son terme général tend vers 0.

4. (2 pts)
   On peut écrire :

```
function x = GPO (A,b,x0,N)
 x=x0
 for i=1:N
   d=A*x-b;
   rho=(d'*d)/((A*d)'*d)
   x=x-rho*d
 end
 endfunction
```

**Exercice 3**

1. (2 pts)
   On a par exemple

```
 ans  =

   1.07   2.01  3.03
   2.06  0.08   2.06
```

   ou

```
 ans  =

    1.01   2.05   3.06
    2.01   0.05   2.02
```

2. (2 pts) La mutation effectuée n'est pas de moyenne nulle autour du point courant. On peut corriger cela :

```
Amut(i,:)=A(i,:)+s*(2*rand(1,n)-1)
```

   ou passer à une mutaition gaussienne :

```
Amut(i,:)=A(i,:)+s*grand(1,1,'nor',0,1)
```

3. (1 pt)
   `s` est le pas de la mutation (dans $\mathbb{R}_+$) et `p` la probabilté de mutation (dans $[0, 1]$).
4. (2 pts) La règle du 1/5 adapte la valeur de $s$ en fonction du taux de succès des mutations dans la génération précédente. On peut aussi adapter le paramètre  en fonction du succès et de la position des succès des mutations précédentes (CSA-ES ou CMA-ES).