4.5 ~~4.4.3~~ Méthodes numériques pour contraintes en égalité

4.5.1 Méthode de pénalisation

On a vu au 1e chapitre que tout problème de minimisation posé sur une partie $U \subset \mathbb{R}^n$

$$\min_{x \in U} f(x)$$

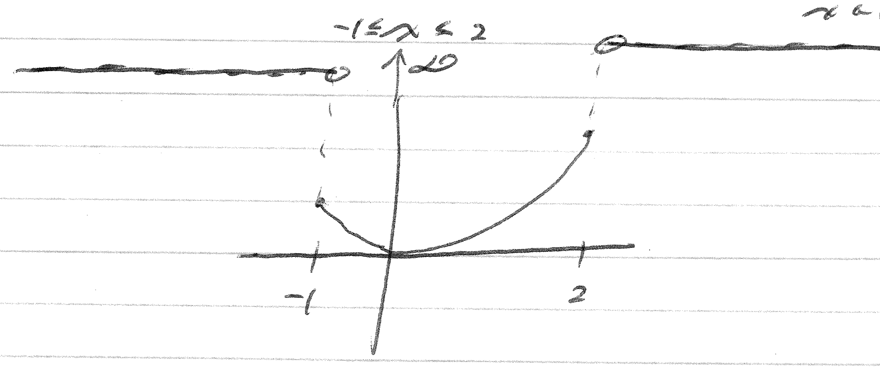
peut s'écrire sous la forme

$$\min_{x \in \mathbb{R}^n} f(x) + I_U(x)$$

où I_U est la fonction indicatrice de l'ensemble U

$$I_U(x) = \begin{cases} 0 & x \in U \\ \infty & \text{sinon} \end{cases}$$

Exemple: $\displaystyle \min_{-1 \le x \le 2} x^2 = \min_{x \in \mathbb{R}} x^2 + I_{[-1,2]}(x)$



Il est logique de vouloir approcher la fonction indicatrice I_U par une fonction plus "régulière".

$$I_U(x) \approx \Psi_r(x)$$

où r est un paramètre de pénalisation $r>0$ et grand. $r\to\infty$

La fonction $\psi_r(x)$ doit vérifié les propriétés.

1) $\psi_r(x) \geq 0 \quad \forall x \in \mathbb{R}^n \quad \forall r>0$

2) $\psi_r(x) = 0 \iff x \in U \quad \forall r>0$

3) $\lim_{r\to\infty} \psi_r(x) = I_U(x)$

Retournons à la contrainte en égalité

$$U = \{x \in \mathbb{R}^n \ / \ g(x) = 0\}$$

où $g: \mathbb{R}^n \to \mathbb{R}^m$

On pose:

$$\psi_r(x) = \frac{r}{2} \| g(x) \|^2$$

Par conséquent, le problème pénalisé s'écrit

$$\min_{x \in \mathbb{R}^n} f(x) + \frac{r}{2} \| g(x) \|^2$$

C'est un problème sans contrainte de fonction

$$\tilde{f}_r(x) = f(x) + \frac{r}{2} \| g(x) \|^2$$

La condition d'optimalité est:

$$\nabla \tilde{f}_r(x) = 0$$

Remarques:

- On peut montrer, sous des conditions assez générales, que la solution x_r (qui va dépendre du paramètre r) converge vers la vraie solution $\bar{x}$.

$$x_r \xrightarrow[r\to\infty]{} \bar{x}$$

Exemple 1 :

$$\min_{y=x^2} \frac{1}{2}\left[(x-1)^2 + (y+1)^2\right]$$

La fonction pénalisée s'écrit :

$$\tilde{f}_r(x,y) = \frac{1}{2}\left[(x-1)^2+(y+1)^2\right] + \frac{r}{2}\left(x^2-y\right)^2$$

La cond. d'opt. est :

$$\begin{cases} x-1 + r(x^2-y)\,2x = 0 \\ y+1 + r(x^2-y)(-1) = 0 \end{cases}$$

$$\Rightarrow \cancel{y+1 = r(x^2-y) \Rightarrow x^2-y = \frac{y+1}{r}}$$

$$\cancel{\Rightarrow x-1 + (y+1)\,2x}$$

Il faut utiliser l'algo. de Newton pour résoudre ce système.

Calcul de la solution exacte :

Pour $y=x^2$

$$f(x,y) = \frac{1}{2}\left[(x-1)^2+(y+1)^2\right] = \frac{1}{2}\underbrace{\left[(x-1)^2+(x^2+1)^2\right]}_{h(x)}$$

$$\frac{dh}{dx} = x-1 + (x^2+1)\,2x = 2x^3+3x-1 = 0$$

qui admet une racine entre 0 et 1

$$\Rightarrow (\bar{x},\bar{y}) = (.313\,,\ .098)$$

Exemple 2: $\min_{Bx=0} \cancel{\frac{1}{2}\|Bx\|} f(x)$

$\Leftrightarrow \min \underbrace{f(x) + \frac{r}{2}\|Bx\|^2}_{\tilde{f}(x)}$

Calcul de $\nabla \tilde{f}(x)$:

$$\nabla \tilde{f}(x) = \nabla f(x) + r B^t B x$$

Si $f(x) = \frac{1}{2}(Ax, x) - (b, x)$

$\Rightarrow \nabla f(x) = Ax - b$

La cond. d'opt. $\nabla \tilde{f} = 0$ s'écrit:

$$Ax - b + rB^t Bx = 0$$

$$\Leftrightarrow \left(A + rB^t B\right) x = b$$

$\downarrow$

matrice dite augmentée

Exemple 3 $\min_{Bx=c} f(x)$

$$\tilde{f}(x) = f(x) + \frac{r}{2}\|Bx - c\|^2$$

$$\Rightarrow \nabla \tilde{f} = \nabla f(x) + r B^t(Bx - c) = 0$$

Si $\nabla f(x) = Ax - b$

$$\Rightarrow Ax - b + rB^t(Bx - c) = 0$$

$$\Leftrightarrow (A + rB^t B)x = b + rB^t c$$

Remarques

- dans la pratique, il est difficile d'obtenir la convergence de l'alg. de Newton pour r grand ($r \to \infty$)
- Pour pallier à cette difficulté, on augmente tranquillement le "r"

On résout une suite de problème

$$\min_{x \in \mathbb{R}^n} f(x) + r_k \frac{\|g(x)\|^2}{2}$$

avec $r_k \to \infty$

4.5.2 Méthode des multiplicateurs de Lagrange

On recherche le pt. minimisant du problème

$$\min_{g(x)=0} f(x) \qquad g : \mathbb{R}^n \to \mathbb{R}^m$$

en résolvant le système non linéaire

$$\begin{cases} \nabla f(x) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x) = 0 \\ g(x) = 0 \end{cases}$$

par l'algo. de Newton.

Remarques:

- il n'est pas facile de choisir de bons pts. de départ (x_0, λ_0) pour cet algo.

- la matrice "tangente" de l'algo. de Newton aura la forme

$$\begin{array}{c|c} & x \quad\quad \lambda \\ \begin{array}{c} x \\ \lambda \end{array} & \left(\begin{array}{c|c} & \\ \hline & 0 \end{array}\right) \end{array}$$

La présence d'un bloc 0 peut causer des problèmes lors de la résolution du système linéaire (pivot nul).

À cause de ces remarques, il est préférable de découpler le calcul de x et de λ. On utilise l'algo. suivant.

<u>Algorithme d'Uzawa:</u>

Pour λ_0 donné,

- Calculer x_k solution de:

$$\min_{x \in \mathbb{R}^n} f(x) + (\lambda_k, g(x)) = 0$$

- Mettre à jour λ_k

$$\lambda_{k+1} = \lambda_k + \rho_k\, g(x_k) \qquad \text{où } \rho_k > 0$$

Remarques:

- si $\rho_k = \rho$, l'algo. est à pas constant.
- la mise à jour de λ est en fait un algo. de type gradient.
- il faut choisir $\rho_k = \rho$ le plus grand possible sans perdre la convergence.

- Sous des conditions générales, l'algo. converge si ρ est suffisamment petit
- En général, cet algo. est lent à converger mais robuste.

Exemple : $$\min_{Bx=c} \frac{1}{2}(Ax,x) - (b,x)$$

$$\nabla f(x) = Ax - b$$
$$g(x) = Bx - c$$

L'algo. d'Uzawa s'écrit:

1^{re} étape:

$$Ax_k - b + B^t\lambda_k = 0$$

$$\Leftrightarrow \quad Ax_k = b - B^t\lambda_k$$

2^e étape:

$$\lambda_{k+1} = \lambda_k + \rho_k(Bx_k - c)$$

Pour accélérer l'algorithme d'Uzawa, il est préférable de combiner à la fois la méthode de pénalisation et celle d'Uzawa. Le principe de base est le suivant

Le problème $\min_{g(x)=0} f(x)$

est équivalent à

$$\min_{g(x)=0} f(x) + \frac{r}{2}\|g(x)\|^2$$

Par conséquent, on peut utiliser la méthode des multiplicateurs de Lagrange pour résoudre ce dernier problème. Pour ~~le~~ cela, on écrit le Lagrangien

$$L_r(x, \lambda) = f(x) + \frac{r}{2}\|g(x)\|^2 + (\lambda, g(x))$$

et on applique la méthode d'Uzawa pour le calcul de x et λ. Cette technique porte le nom de méthode du Lagrangien Augmenté

Algorithme du Lagrangien Augmenté

Pour λ_0 donné

1e étape: calculer x_k en résolvant le problème de minimisation sans contrainte

$$\min_{x \in \mathbb{R}} L_r(x, \lambda_k)$$

2e étape: mise à jour de λ

$$\lambda_{k+1} = \lambda_k + r\, g(x_k)$$

Remarques:

- le critère d'arrêt peut être

$$\|g(x_k)\| < \epsilon \quad \text{ou} \quad \|\lambda_{k+1} - \lambda_k\| < \epsilon$$

- il n'est pas nécessaire de prendre r grand car l'algo. converge très bien.