PROGRAMMAZIONE NON LINEARE

by Mattia Campolese – www.matsoftware.it - 28/01/2008 – Ricerca Operativa 1

NON VINCOLATA

Problemi di minimo di funzioni f(x) con $x \in \mathbb{R}^n$.

Condizioni:

- **1° ordine:**
  - *necessarie*: se $f(x) \in \mathbb{C}_1$ e x* è min. locale ⇒ $\nabla f(x^*) = 0$
  - *sufficienti*: se $f(x) \in \mathbb{C}_1$ ed è convessa allora x* è min. globale ⇔ $\nabla f(x^*) = 0$
- **2° ordine:**
  - *necessarie*: se $f(x) \in \mathbb{C}_2$ e x* è min. locale ⇒ $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ è semidef. positiva
  - *sufficienti*: se $f(x) \in \mathbb{C}_2$ e $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ è semidef. positiva ⇒ x* è min. locale

Condizioni di esistenza di un minimo globale:

- $f(x) \in \mathbb{C}_1$ su X chiuso e limitato (th. Weirstrass)
- $f(x) \in$ insieme di livello α chiuso e limitato; sup. di livello = $\{ x \in \mathbb{R}_n \,/\, f(x) \leq \alpha \}$
- f(x) è coerciva ($\lim_{x \to +\infty} f(x) = +\infty$)

Criteri per gli algoritmi:

- **Condizione d'angolo:**
  La direzione di discesa ed il gradiente devono formare un angolo maggiore di 90°.
  $d_k^T \cdot \nabla f(x_k) \leq -\epsilon \left| d_k^T \right| \cdot \left| \nabla f(x_k) \right|$

- **Condizione di sufficiente riduzione:**
  La diminuzione della funzione deve essere significativa.
  Ponendo $f(x_k + \alpha d_k) = \varphi(\alpha)$ con $\alpha > 0$, si ha che la funzione localmente è compresa tra $\varphi(0) = f(x_k)$ e $\varphi(\alpha) = \alpha\varphi'(\alpha) + \varphi(0)$. Si impone che sia inferiore alla loro combinazione conica:
  $f(x_k + \alpha d_k) \leq f(x_k) + \gamma \alpha \nabla f(x_k)^T d_k$ con $\gamma \in [0,1]$

- **Condizione di Wolfe:**
  La pendenza della funzione deve diminuire, deve tendere al gradiente nullo.
  $\nabla f(x_k + \alpha d_k)^T d_k \geq c \nabla f(x_k)^T d_k$ con $c \in [\gamma, 1]$

Convergenza degli algoritmi:

- **Globale**:
  $\lim_{k \to +\infty} \nabla f(x^k) = 0$

- **Locale**:
  $\lim_{k \to +\infty} \bar{x}$ nell'intorno del punto stazionario

- **Rapidità di convergenza**:
  $\lim_{k\to+\infty}\frac{|x^{k+1}-\overline{x}|}{|x^k-\overline{x}|}=c$
  - se c = ∞ l'algoritmo *non converge;*
  - se c ≥ 1 ha convergenza *sublineare;*
  - se 0 > c > 1 ha convergenza *lineare;*
  - se c = 0 ha convergenza *superlineare;*
  - se $\lim_{k\to+\infty}\frac{|x^{k+1}-\overline{x}|}{|x^k-\overline{x}|^2}=c$ ha convergenza *quadratica;*

Algoritmi di discesa:

- **Line search esatta:**

  - direzione scelta con il metodo del gradiente: $d_k=-\nabla f(x_k)$
  - passo α scelto minimizzando $f(x_k + \alpha d_k) = \varphi(\alpha)$

  Schema:

```
Si pone k:=0 e si fissa x0 ∈ ℝn;
while ∇f(xk) ≠ 0 do
     {
      dk := - ∇f(xk);
      αk := min[f(xk + αkdk) ];
      xk+1 := xk + αkdk;
      k := k + 1;
     }
```

- **Line search approssimata:**
  Rispetto alla precedente cambia la modalità di scelta del passo:

  <u>*Metodo di Armijo (o backtracking):*</u>

  Scelto un α_0 iniziale suff. grande, il passo è determinato calcolando $\alpha^{i+1}=\sigma\alpha^i$ con σ ∈ [0,1/2] detto *fattore di contrazione* fino a quando il passo non soddisfa la condizione di sufficiente riduzione.

  Schema:

```
...
Si pone i:=0 e si fissa α0 ∈ ℝ;
while f(xk + αidk) > f(xk)) + γαi ∇f(xk)T dk do
     {
      αi+1 := σ αi;
      i := i + 1;
     }
```

  <u>*Metodo di bisezione:*</u>

  Identico ad Armijo con la condizione aggiuntiva di rispetto del criterio di Wolfe. Nel calcolo del passo infatti se il precedente non soddisfa la suff. rid. e il seguente non soddisfa Wolfe, si può utilizzare un passo intermedio tra i due: $\alpha^{i+1}=\frac{\alpha^{i}+\alpha^{i-1}}{2}$. Nel caso la condizione non venga ancora verificata, si può ripetere. Vantaggio è la garanzia della convergenza globale, svantaggio è passaggio aggiuntivo.

<u>*Metodo della interpolazione:*</u>

Rappresentazione approssimata della φ(α) con un polinomio quadratico, si minimizza e si trova l' α che rispetta la suff. riduzione $\phi(\alpha)=a\alpha^2+b\alpha+c$. Riduzione più forte.

Essendo

$$c=\phi(0)=f(x_k)$$

$$b=\phi'(0)=\nabla f(x_k)^T\cdot d_k$$

$$a=\frac{\phi(\alpha_i)-b\alpha_i-c}{\alpha_i^2}$$

il valore interpolato sarà:

$$\alpha_{i+1}=min\left(\frac{\phi(\alpha_i)-b\alpha_i-c}{\alpha_i^2}*\alpha_{i+1}^2+\nabla f(x_k)^T\cdot d_k*\alpha_{i+1}+f(x_k)\right) \text{ ovvero}$$

$$\alpha^{i+1}=-\frac{\nabla f(x_k)^T\cdot d_k*\alpha_i^2}{2\left[-f(x_k)-\alpha_i\nabla f(x_k)^T\cdot d_k+f(x_k+\alpha_i d_k)\right]}$$

Schema:

```
...
Si pone i:=0 e si fissa α₀ ∈ ℝ;
while f(xₖ + αᵢdₖ) > f(xₖ)) + γ αᵢ ∇f(xₖ)ᵀ dₖ do
    {
```

$$\alpha^{i+1}=-\frac{\nabla f(x_k)^T\cdot d_k*\alpha_i^2}{2\left[-f(x_k)-\alpha_i\nabla f(x_k)^T\cdot d_k+f(x_k+\alpha_i d_k)\right]}$$

```
     i := i + 1;
    }
```

- **Metodo di Newton puro**

  Si trova insieme il passo e la direzione, approssimando la funzione con una forma quadratica, ricavando il vettore direzione dallo sviluppo in serie di Taylor e applicando la condizione necessaria del 2° ordine.

  Convergenza quadratica al posto di quella lineare del metodo del gradiente.

  $$f(x_k+h)=f(x_k)+\nabla f(x_k)^T\cdot h+\frac{1}{2}h^T\nabla f(x_k)h+\beta(h,x_k) \text{ il cui minimo si ha in}$$

  $$\bar{h}=-\frac{\nabla f(x_k)}{\nabla^2 f(x_k)^{-1}}$$

  Schema:

```
Si pone k:=0 e si fissa x₀ ∈ ℝₙ;
while ∇f(xₖ) ≠ 0 do
    {
     h* := - ∇²f(xₖ)⁻¹ * ∇f(xₖ);
     xₖ₊₁ := xₖ + h*;
     k := k + 1;
    }
```

- **Metodo di Newton modificato**
  Globalmente convergente e localmente con rapidità quadratica.
  Se non è possibile o non conviene seguire la direzione di Newton, si effettua una line search esatta altrimenti si applica Newton tradizionale.
  Schema:

```
Si pone k:=0 e si fissa x₀ ∈ ℝₙ;
while ∇f(xₖ) ≠ 0 do
     {
     h* := - ∇²f(xₖ)⁻¹ * ∇f(xₖ);
       se ∇f(xₖ)ᵀ * h* < 0 allora
          {
           line search con dₖ := h*
          }
       altrimenti
          {
           line search con dₖ := - ∇f(xₖ)
          }
     k := k + 1;
     }
```

VINCOLATA

Problemi di minimo di funzioni f(x) con $x \in X \subset \mathbb{R}^n$ caratterizzato come

$X = \{x \in \mathbb{R}^n / h(x) = 0 ; g(x) \geq 0\}$

Esistenza minimo globale

Il minimo globale esiste se

- la funzione è coerciva
- $f(x) \in \mathbb{C}_1$ su X chiuso e limitato

se esso esiste allora sarà o

- un punto di non regolarità
- un punto KKT

Condizioni di regolarità

I punti di regolarità sono quelli per cui lo jacobiano dei vincoli attivi ha rango pieno.

Condizioni di Karush Kuhn Tucker (KKT)

Si considera la funzione lagrangiana $L(x, \lambda, \mu) = f(x) - \lambda^T h(x) - \mu^T g(x)$. Si impone:

$\nabla_x L(\bar{x}, \lambda) = 0$
$\mu^T g(\bar{x}) = 0$
$\mu \geq 0$
ovvero per ogni combinazione di vincoli attivi si determinano i punti candidati al minimo.

Algoritmi per PNL vincolata

Vengono guidati gli algoritmi al rispetto dei vincoli attivi per evitare di processarne 2^k combinazioni possibili.

- **Funzioni di penalità**
  Usate per comodità solo con vincoli di uguaglianza. La funzione obiettivo diventa una $F(x)=f(x)+p(x)$ con p(x) nullo nell'insieme ammissibile e crescente fuori.

- **Funzioni di barriera**
  Usate con i vincoli di diseguaglianza. Usata spesso la forma logaritmica, sono tali per cui la funzione cresce all'avvicinarsi del vincolo attivo.