Metodi Matematici per l'Ingegneria - 08/07/2013 - Parte A

- **Esercizio: 1-OT** Si consideri la funzione $f(\mathbf{x} = (x, y)) = a\ x^2 + b\ y^2 + c\ x\ y + d\ y$, con $a, b, c, d \in \mathbb{R}$:

  1. trovare i punti stazionari;
  2. scrivere la condizione sufficiente per un punto di massimo;
  3. posto $a = -1$ sia $\Omega = \{(c, b) \in \mathbb{R}^2$: il punto stazionario risulti un punto di massimo$\}$, disegnare Ω nel piano $\{c, b\}$;
  4. verificare che per $a = -1$, $b = -1$, $c = \sqrt{2}$, $d = 2$ il punto stazionario $\mathbf{x}^*$ della funzione $F(\mathbf{x} = (x, y)) = a\ x^2 + b\ y^2 + c\ x\ y + d\ y + e$, con $e \in \mathbb{R}$ risulta massimo e calcolare il valore $F(\mathbf{x}^*)$.
  5. Con l'utilizzo della function **fminunc** trovare i massimi della funzione alla funzione

     $$\Phi(\mathbf{x}) = [F(\mathbf{x})]^3 \quad \text{per} \quad a = -1, \quad b = -1, \quad c = \sqrt{2}, \quad d = 2$$

     applicando il metodo Quasi-Newton con le seguenti **options**:
     - il tipo di problema non è LargeScale;
     - si sceglie l'update *bfgs* ;
     - la direzione iniziale sia determinata prendendo come matrice iniziale approssimante l'Hessiana, la matrice di identità;
     - si fornisce il gradiente;
     - si assegnano le seguenti tolleranze: TolFun:1.e-14; TolX: 1.d-14.
  6. Nei casi

     $$e = -1\,, \qquad e = -4\,,$$

     si esegua il metodo Quasi-Newton considerando i seguenti punti di iniziali:

     $$\mathbf{x}_0 = (2, -1.5)\,, \qquad \mathbf{x}_0 = (2, 1).$$

     Riportare per le iterate {0,1, e l'ultima } i seguenti dati:

     iteration   Func-count   f(x)   first-order condition

     ed il punto di massimo ottenuto.

     Confrontare i risultati ottenuti nel caso $e = -1$ rispetto al caso $e = -4$, sia in termini del valore del punto di massimo che del valore massimo dell'obbiettivo Φ; si guardi nei due casi anche il grafico di Φ sul dominio $[0.5, 1.6] \times [1, 2.3]$ nei due precedenti casi.
  7. Si fornisca una motivazione del diverso comportamento.
  8. Considerando il caso $e = -1$ quale velocità di convergenza si osserva per i due punti iniziali $\mathbf{x}_0 = (2, -1.5)$ e $\mathbf{x}_0 = (2, 1)$.
  9. Considerando il caso $e = -4$ quale velocità di convergenza si osserva per i due punti iniziali $\mathbf{x}_0 = (2, -1.5)$ e $\mathbf{x}_0 = (2, 1)$.

1-TD. Siano $g(x) = x^2 \sin(x)$ ed $f(x) = x^2 \sin(9x)$. Tracciare un grafico qualitativo di $\mathrm{Re}(c_n)$, $\mathrm{Im}(c_n)$ e $|c_n|$ per g ed f. Commentare le differenze.

2-TD. Sia $f : (0, 2\pi) \to \mathbb{R}$ data da $f(t) = -t - 1$. Siano c_n i coefficienti di Fourier di f. Tracciare un grafico qualitativo di $\mathrm{Re}(c_n)$, $\mathrm{Im}(c_n)$ e $|c_n|$. Si consideri la succesione di polinomi trigonometrici (serie di Fourier troncata)

$$S_k(t) = \sum_{n=-k}^{k} c_n e^{int}, \qquad \text{per } k \in \mathbb{N}.$$

Tracciare un grafico qualitativo di S_k per alcuni valori di k e commentare la convergenza.

Metodi Matematici per l'Ingegneria - 08/07/2013 - Parte B

- **Esercizio: 1-OT**

  1. Sia $\mathbf{s}_k$ una di direzione di discesa, si scriva la regola di Armijo per la ricerca del passo con percentuale ρ. Posto $\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{s}_k)$ si fornisca una interpretazione geometrica della regola di Armijo con parametro ρ. Si consideri congiuntamente la seguente regola di ammissibilità del passo :

  $$f(\mathbf{x}_k) + \sigma\ \alpha\ \nabla f(\mathbf{x}_k)^T \mathbf{s}_k \leq f(\mathbf{x}_k + \alpha \mathbf{s}_k).$$

  Scelto $0 < \ \rho < \ \sigma < 1$ si fornisca una interpretazione geometrica dell'insieme dei passi ammissibili α soddisfacenti contemporaneamente le due regole.

- **Esercizio: 2-OT**

  1. Sia $\mathbf{x}_k$ convergente a $\mathbf{x}^*$ fornire la definizione di convergenza superlineare e quadratica.
  2. Si definisca lo step di Newton. Formulare le proprietá di convergenza del metodo di Newton. Sotto quali condizioni è garantita la massima velocità di convergenza.
  3. Sia $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x} - b^T\mathbf{x}$ con $\mathbf{A}$ matrice simmetrica e definita positiva . Data l'iterata $\mathbf{x}_k$ determinare lo step δ_k di Newton, verificare che δ_k risulta una direzione di discesa e calcolare $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{opt}\delta_k$ con $\alpha_{opt} = argmin\{f(\mathbf{x}_k + \alpha\delta_k), \alpha > 0\}$. Commentare il risultato trovato.

  oppure

- **Esercizio: 2-OT** Sia $\mathbf{s}_k \in \mathbb{R}^n$ una direzione di discesa in $\mathbf{x}_k \in \mathbb{R}^n$ per la funzione obiettivo $f(\mathbf{x})$. Posto $\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{s}_k)$ si scelga il passo ottimale α_{ott} per cui $\phi(\alpha_{ott}) = argmin\{\phi(\alpha), \alpha > 0\}$,

  - definito $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{ott}\mathbf{s}_k$ quale proprietà verifica $\nabla f(\mathbf{x}_{k+1})$?
  - se $f(\mathbf{x}) = = \frac{1}{2}\ \mathbf{x}^T\mathbf{H}\mathbf{x} - c^T\mathbf{x}$, con $\mathbf{H}$ simmetrica e definita positiva e $\mathbf{c} \in \mathbb{R}^n$ si calcoli il passo ottimale α_{ott} relativo alla direzione di discesa $\mathbf{s}_k$;
  - si consideri il seguente algoritmo:
    - ∗ dati $(\mathbf{x}_0,\ \mathbf{H},\ \mathbf{c})$
    - ∗ $\mathbf{r}_0 = \mathbf{H}\mathbf{x}_0 - \mathbf{c}$
    - ∗ for k=0,1,2,...
    - ∗ $\mathbf{m}_k = \mathbf{H}\mathbf{r}_k$
    - ∗ $\alpha_k = \frac{\mathbf{r}_k^T\mathbf{r}_k}{\mathbf{r}_k^T\mathbf{m}_k}$
    - ∗ $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k\mathbf{r}_k$
    - ∗ $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{m}_k$
    - ∗ end

    si dimostri che l'algoritmo è una formulazione equivalente di un ben noto metodo di minimizzazione;
  - quali sono le proprietà di convergenza del precedente algoritmo?

1-TD. Siano $g : [0, 2\pi] \to \mathbb{R}$ ed $f : [0, 2\pi] \to \mathbb{R}$ definite da $g(x) = |x|$ ed $f(x) = |x - \pi|$. Stabilire (per g e per f) se è corretto aspettarsi che $\mathrm{Re}(c_n) = 0$ per ogni $n \neq 0$ o che $\mathrm{Im}(c_n) = 0$ per ogni n.

2-TD. Sia $f : [0, 2\pi] \to \mathbb{R}$ e sia $\Delta t = 2\pi/N$. Sia y il vettore (di campionamento) con componenti $y_n = f(n\Delta t)$ per $n = 0, ..., N-1$. Sia $Y = \mathrm{fft}(y)/N = \mathbb{F}y/N$ il vettore MATLAB della trasformata discreta, dove $\mathbb{F}$ è la matrice con elementi $\mathbb{F}_{nm} = w^{-nm}$ per $w = e^{i\Delta t}$. Quali delle seguenti proprietà sono verificate dalla $\mathbb{F}$?

1. $\mathbb{F}$ è simmetrica,
2. $\mathbb{F}$ è una matrice reale $N \times N$,
3. $\mathbb{F}$ è hermitiana (o unitaria),
4. $\mathbb{F}$ è diagonale,
5. $\mathbb{F}$ è invertibile.

- **Esercizio: 1-OT** Si consideri la funzione $f(\mathbf{x} = (x,y)) = a\ x^2 + b\ y^2 + c\ x\ y + d\ y$, con $a,b,c,d \in R$:

  1. trovare i punti stazionari.
     $\nabla f(x,y) = (\ 2\ a\ x + c\ y,\ 2\ b\ y + c\ x + d\ )^T = 0$ $[2*a*x + c*y = 0,\quad x = -c*y/(2*a)$ da cui $y(2*b - c^2/(2*a)) = -d\ \ y = -2ad/det$ con $det = 4*a*b - c^2]$ per cui $\mathbf{x}^* = (\frac{cd}{4ab-c^2}, \frac{-2\ a\ d}{4\ a\ b-c^2}$ è l'unico punto stazionario.
  2. scrivere la condizione sufficiente per i massimi.
     $\nabla f(\mathbf{x}^*) = 0$ e $Hf(\mathbf{x}^*)$ definita negativa
  3. Posto a=-1 trovare disegnare l'insieme $\Omega = \{(b,c) \in R^2\}$ affinché il punto stazionario risulti un punto di massimo
     $Hf(\mathbf{x}) = \begin{pmatrix} 2\ a & c \\ c & 2\ b \end{pmatrix}$ impongo che sia definita negativa:
     $tr(Hf(\mathbf{x}^*)) = 2(-1+b) < 0,\quad det(Hf(\mathbf{x}^*)\ ) = 4ab - c^2 = -4b - c^2 > 0$ da cui $b < 1$, e $b < -c^2/4$ quindi $\Omega = \{(b,c) \in R^2 : b < -\frac{c^2}{4}\}$.
  4. Verificare che per a=-1, b=-1, c=$\sqrt{2}, d = 2$ il punto stazionario $\mathbf{x}^*$ della funzione $F(\mathbf{x} = (x,y)) = a\ x^2 + b\ y^2 + c\ x\ y + d\ y + e$, con $e \in R$ risulta massimo e calcolare il valore $F(\mathbf{x}^*)$.
  5. Utilizzando la function **fminunc** per trovare i massimi della funzione
     $\Phi(\mathbf{x}) = [\ F(\mathbf{x})]^3$ per $a = -1,\ b = -1,\ c = \sqrt{2},\ d = 2$ applicando il metodo Quasi-Newton.

     Si assegni nelle **options** che :
     - il tipo di problema non è LargeScale;
     - si sceglie l'update *bfgs*
     - la direzione iniziale sia determinata prendendo come matrice iniziale approssimante l'Hessiana, la matrice di identità;
     - si fornisce il gradiente
     - si assegnano le seguenti tolleranze: TolFun:1.e-14; TolX: 1.d-14;

     Riportare per le iterate {0,1, e l'ultima } i seguenti dati:
     Metodo **BFGS**: iteration Func-count f(x) first-order condition
     ed il punto di massimo ottenuto.

     - options = optimset('LargeScale','off','MaxFunEvals',250);
     - options = optimset(options,'Display','iter','TolFun',1e-10,'TolX',1e-8,'GradObj','on');
     - options=optimset(options,'HessUpdate','bfgs','InitialHessType','scaled-identity');
     - disp(' Quasi -Newton : BFGS con gradiente ')
     - [x,fval,exitflag,output] = fminunc(@pbnv-objfun,x0,options)
     - costruisco il file pbnv_objfun.m che contiene l'obiettivo ed il suo gradiente
       - function [phi,grad] = pbnv_objfun(x);
       - a=-1;b=2;c=$\sqrt{2}$;
       - d=2;e=-1;
       - f= $a*x(1)^2 + b*x(2)^2 + c*x(1)*x(2) + d*x(2)$;
       - grad(1)= -$(3*f^2*(2*a*x(1) + c*x(2)))$;
       - grad(2)= - $(3*f^2*(2*b*x(2) + c*x(1) + d))$;
       - phi= - f^3;
       - end

Il massimo di Φ è raggiunto in $\mathbf{x}^* = (,)^T$ che coincide con il minimo della funzione dell'esercizio **Esercizio: 2-OT**.

Dal calcolo esplicito dell'Hessiana nel punto di minimo si giustifichi l'efficienza del emtodo per j=1 e j=2, in termini di velocità di convergenza. verificare che il punto stazionario di $f(\mathbf{x})$ risulta punto di minimo per $\Phi(\mathbf{x})$

Essendo $\mathbf{H}$ definita negativa il punto stazionario $\mathbf{x}^*$ per f è punto di massimo, inoltre $f(\mathbf{x}^*) = \frac{1}{2}\ (\mathbf{x}^*)^T\mathbf{H}\mathbf{x}^* - (\mathbf{H}\mathbf{x}^*)^T\mathbf{x}* = -\frac{1}{2}\ (\mathbf{x}^*)^T\mathbf{H}\mathbf{x}^* > 0$

L' Hessiana di Φ è data da: $\nabla^2\phi(\mathbf{x}) = f(\mathbf{x})\mathbf{H} + \nabla f(\mathbf{x})(\mathbf{H}\mathbf{x} - \mathbf{c})^T = f(\mathbf{x})\mathbf{H} + (\mathbf{H}\mathbf{x} - \mathbf{c})(\mathbf{H}\mathbf{x} - \mathbf{c})^T$ da cui $\nabla^2\phi(\mathbf{x}^*) = f(\mathbf{x}^*)\mathbf{H} + ||\mathbf{H}\mathbf{x}^* - \mathbf{c}||_2^2 = f(\mathbf{x}^*)\mathbf{H}$ ed essendo $f(\mathbf{x}^*) > 0$ allora $\nabla^2\Phi(\mathbf{x}^*)$ risulta definita positiva e quindi $\mathbf{x}^*$ è punto di minimo per ϕ.

Il gradiente di Φ risulta $\nabla\Phi(\mathbf{x}) = 2f(\mathbf{x})\nabla f(\mathbf{x})$, quindi $\mathbf{x}^*$ è anche l'unico punto stazionario per Φ essendo $f(\mathbf{x}) \geq f(\mathbf{x}^*) \geq 0\forall\mathbf{x}$ e per d=1 $f(\mathbf{x}^*) = 0$ e per d=2 $f(\mathbf{x}^*) = 2$.

Inoltre si ha $\nabla^2\Phi(\mathbf{x}) = 2\ \nabla f(\mathbf{x})\nabla f(\mathbf{x})^T\ + 2\ f(\mathbf{x})\nabla^2 f(\mathbf{x})$.

Per j=1 si ha $\mathbf{x}^*:\ f(\mathbf{x}^*) = 0$ ed allora $\nabla^2\Phi(\mathbf{x}^*) = \mathbf{0}$. Perciò $\nabla^2 f(\mathbf{x}^*)$ é singolare ed allora la convergenza del metodo Quasi-Newton degenera risultando solo di tipo lineare.

Mentre per d=2 si ha $f(\mathbf{x}^*) = 2$ ed allora

$$\nabla^2\Phi(\mathbf{x}^*) = 2\ \{2\ \nabla^2 f(\mathbf{x}^*)\} = 4\begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}.$$

matrice definita positiva. La convergenza del metodo Quasi-Newton risulta superlineare.

- **Esercizio: 1-OT**

  1. Sia $\mathbf{s}_k$ una di direzione di discesa, si scriva la regola di Armijo per la ricerca del passo con percentuale ρ. Posto $\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{s}_k)$ si fornisca una interpretazione geometrica della regola di Armijo con parametro ρ. Si consideri congiuntamente la seguente regola di ammissibilità del passo :

  $$f(\mathbf{x}_k) + \sigma \ \alpha \ \nabla f(\mathbf{x}_k)^T \mathbf{s}_k \leq f(\mathbf{x}_k + \alpha \mathbf{s}_k)$$

  Scelto $0 < \ \rho < \ \sigma < 1$ si fornisca una interpretazione geometrica dell'insieme dei passi ammissibili α soddisfacenti contemporaneamente le due regole.

- **Esercizio: 2-OT**

  1. Sia $\mathbf{x}_k$ convergente a $\mathbf{x}^*$ fornire la definizione di convergenza superlineare e quadratica.
     - Sia $\mathbf{x}_k$ convergente a $\mathbf{x}^*$ fornire la definizione di convergenza superlineare :

       $$||\mathbf{x}_{k+1} - \mathbf{x}^*|| \leq \rho_k \ ||\mathbf{x}_k - \mathbf{x}^*||, \text{ with } \rho_k \to 0 \ \text{ per} k \to \ \infty$$

     - Sia $\mathbf{x}_k$ convergente a $\mathbf{x}^*$ fornire la definizione di convergenza quadratica:

       $$\text{esiste C } > 0 : \ \ ||\mathbf{x}_{k+1} - \mathbf{x}^*|| \leq C \ ||\mathbf{x}_k - \mathbf{x}^*|| \ ^2, \ \forall \ k \geq k_0$$

  2. Formulare le proprietá di convergenza del metodo di Newton. Sotto quali condizioni è garantita la massima velocità di convergenza Lo step di Newton è definito da

  $$\left\{ \begin{array}{l} \nabla^2 f(\mathbf{x}^{(k)}) \ \delta^{(k)} = -\nabla f(\mathbf{x}^{(k)}) \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta^{(k)} \end{array} \right\}$$

  Sia $\mathbf{x}^* \ : \ \nabla f(\mathbf{x}^*) = 0$ con $\nabla^2 f(\mathbf{x}^*)$ é definita positiva, quindi $\mathbf{x}^*$ risulta un minimo relativo isolato, inoltre $f \in C^3(B*)$ oppure $f \in Cr(B^*)$ e $\nabla f(\mathbf{x}^*)$ risulta Lipschitziana in $B^* = \mathbf{x} : ||\mathbf{x} - \mathbf{x} * || < \rho$ allora la convergenza risulta quadratica. Se f é solamente in $C^2(B*)$ la convergenza risulta superlineare. Se $\nabla^2 f(\mathbf{x}^*)$ é singolare allora la convergenza é solo di tipo lineare.

  3. Sia $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x} - b^T \mathbf{x}$ con $\mathbf{A}$ matrice simmetrica e definita positiva . Data l'iterata $\mathbf{x}_k$ determinare lo step δ_k di Newton, verificare che δ_k risulta una direzione di discesa e calcolare $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{opt}\mathbf{p}_k$ con $\alpha_{opt} = argmin\{f(\mathbf{x}_k + \alpha \mathbf{p}_k), \alpha > 0\}$. Commentare il risultato trovato.

- **Esercizio: 2-OT** Sia $\mathbf{s}_k \in R^n$ una direzione di discesa in $\mathbf{x}_k \in R^n$ per la funzione obiettivo $f(\mathbf{x})$. Posto $\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{s}_k)$ si scelga il passo ottimale α_{ott} per cui $\phi(\alpha_{ott} 0 = argmin\{\phi(\alpha), \alpha > 0\}$

  - definito $\mathbf{x}k + 1 = \mathbf{x}_k + \alpha_{ott}\mathbf{s}_k$ quale proprietà verifica $\nabla f(\mathbf{x}_{k+1}$?
  - se $f(\mathbf{x}) = = \frac{1}{2} \ \mathbf{x}^T \mathbf{H}\mathbf{x} - c^T \mathbf{x}$, con $\mathbf{H}$ simmetrica e definita positiva e $\mathbf{c} \in R^n$ si calcoli il passo ottimale α_{ott} relativo alla direzione di discesa $\mathbf{s}_k$
  - si consideri il seguente algoritmo:
    - dati $(\mathbf{x}_0, \ \mathbf{H}, \ \mathbf{c})$
    - $\mathbf{r}_0 = \mathbf{H}\mathbf{x}_0 - \mathbf{c}$

* for k=0,1,2,...
* $\mathbf{m}_k = \mathbf{H}\mathbf{r}_k$
* $\alpha_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_k^T \mathbf{m}_k}$
* $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{r}_k$
* $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{m}_k$
* end

si dimostri che l'algoritmo è una formulazione equivalente di un ben noto metodo di minimizzazione.

- quali sono le proprietà di convergenza del precedente algoritmo.