ESAME di METODI DI OTTIMIZZAZIONE PER BIG DATA - Compito A

Cognome : **Nome** :

Esercizio 1. Si consideri il seguente problema:

$$\min_{x \in \Re^2} \left(x_1^2 + 1 + (x_2 - 1)^2 \right)^2$$

(i) Dire se è possibile concludere apriori esistenza e/o unicità della soluzione.

(ii) Individuare i punti critici della funzione e la loro natura.

(iii) Nel punto (1 1) calcolare la direzione dell'antigradiente e quella di Newton (se possibile).

Esercizio 2. Si consideri il problema:

$$\min_{x \in S} \quad \tfrac{1}{2}(x_1^2 + 4x_1x_2 + 3x_2^2) + x_2$$

(i) Assumendo $S = \Re^2$ dire se è possibile concludere apriori esistenza e/o unicità della soluzione ed eventualmente determinarle.

(ii) Si assuma $S = \{x \in \Re^2 : x_1^2 + x_2^2 \leq 9, \ -2x_1 - 3x_2 \leq 6, \ 2x_1 - x_2 \leq 2\}$. In questo caso:

(a) E'possibile concludere apriori l'esistenza della soluzione?

(b) Scrivere le condizioni di KKT e verificare se il punto $(\frac{2}{3} \quad -\frac{2}{3})$ le soddisfa. Quali conclusioni si possono trarre sul punto?

(c) Effettuare un'iterazione del metodo di Frank Wolfe partendo dal punto (1 1), con parametri di Armijo pari a $\Delta = 1$, $\gamma = \delta = \frac{1}{4}$.

Esercizio 3. Sia dato il seguente problema non vincolato:

$$\min \frac{1}{2}\left(6x_1^2 + 6x_1x_2 + 4x_2^2\right)\left(x_3^2 + 1\right) + (x_3 - 2)^2 - 3x_1 - x_2 + 2x_3$$

Effettuare la migliore decomposizione in blocchi possibile e una iterazione a partire dal punto $(-1 \ 1 \ 0)$ di un metodo di decomposizione scelto in base alle proprietà della funzione giustificando la scelta sia della decomposizione che dell'algoritmo. Si illustrino le proprietà di convergenza dell'algoritmo scelto in questo caso.

Esercizio 4. Sia dato il seguente insieme di dati:

$$\mathcal{T} = \left\{ \left(\begin{pmatrix}1\\2\end{pmatrix}, 1\right), \left(\begin{pmatrix}2\\4\end{pmatrix}, 1\right), \left(\begin{pmatrix}3\\2\end{pmatrix}, 1\right), \left(\begin{pmatrix}5\\2\end{pmatrix}, 1\right), \left(\begin{pmatrix}5\\4\end{pmatrix}, 1\right), \right.$$
$$\left. \left(\begin{pmatrix}3\\3\end{pmatrix}, -1\right), \left(\begin{pmatrix}2\\0\end{pmatrix}, -1\right), \left(\begin{pmatrix}1\\-2\end{pmatrix}, -1\right), \left(\begin{pmatrix}5\\-1\end{pmatrix}, -1\right), \left(\begin{pmatrix}6\\1\end{pmatrix}, -1\right) \right\}$$

e si supponga di voler costruire una SVM per risolvere il corrispondente problema di classicazione utilizzando il kernel lineare ($k(x_i, x_j) = x_i^T x_j$).

(i) Scrivere il problema duale che deve essere risolto per addestrare la SVM (con parametro $C = 10$).

(ii) Sapendo che la soluzione del problema duale ‘e la seguente (approssimata alla seconda cifra decimale):

$$\lambda = \begin{pmatrix} 0 \\ 0 \\ 8.92 \\ 10 \\ 0 \\ 10 \\ 1.65 \\ 0 \\ 0 \\ 7.26 \end{pmatrix}$$

individuare i vettori di supporto, calcolare l'iperpiano di separazione ottimo $w^T x + b$ e dire quali valori delle variabili ξ_i possono essere determinati.

(iii) Come classifica la SVM così addestrata il punto $x^T = (-3\ 8)$?

(iv) E' possibile dedurre la separabilità lineare dei due insiemi di dati?Come si calcola l'errore sul training set?

(iv) Addestrando nuovamente la SVM con kernel polinomiale con $\gamma = 0$ e $p = 3$, la soluzione diventa

$$\lambda = \begin{pmatrix} 0.38 \\ 0.0006 \\ 0 \\ 0 \\ 0.062 \\ 0.25 \\ 0.19 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Che cosa si può concludere guardando questa soluzione?(non è necessario eseguire nessun calcolo)

Esercizio 5 Sia dato il seguente problema:

$$\begin{array}{ll} \min & f(\alpha) \\ & -\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 = 0 \\ & 0 \leq \alpha_i \leq 3 \quad i = 1, \ldots, 4 \end{array} \tag{1}$$

Supponiamo che nel punto $\bar{\alpha} = (3\ 1\ 0\ 2)$, il gradiente della funzione valga:

$$\nabla f(\bar{\alpha}) = \begin{pmatrix} 3 \\ \beta \\ -1 \\ \gamma \end{pmatrix}$$

(i) Per quali valori di γ e β il punto $\bar{\alpha}$ è ottimo?

(ii) Per quali valori di γ e β il working set di SVM$_{light}$ è $\{1, 4\}$?

ESAME di METODI DI OTTIMIZZAZIONE PER BIG DATA - Compito B

Cognome : **Nome** :

Esercizio 1. Si consideri il seguente problema:

$$\min_{x\in\Re^2} \left((x_1-3)^2+2x_2^2+2\right)^2$$

(i) Dire se è possibile concludere apriori esistenza e/o unicità della soluzione.

(ii) Individuare i punti critici della funzione e la loro natura.

(iii) Nel punto (1 1) calcolare la direzione dell'antigradiente e quella di Newton (se possibile).

Esercizio 2. Si consideri il problema:

$$\min_{x\in S} \quad \tfrac{1}{2}(2x_1^2+8x_1x_2+3x_2^2)+x_2$$

(i) Assumendo $S=\Re^2$ dire se è possibile concludere apriori esistenza e/o unicità della soluzione ed eventualmente determinarle.

(ii) Si assuma $S=\{x\in\Re^2 : x_1^2+x_2^2\le 9,\ -2x_1-3x_2\le 6,\ 2x_1-x_2\le 2\}$. In questo caso:

(a) E'possibile concludere apriori l'esistenza della soluzione?

(b) Scrivere le condizioni di KKT e verificare se il punto $(\frac{3}{5}\ \ -\frac{4}{5})$ le soddisfa. Quali conclusioni si possono trarre sul punto?

(c) Effettuare un'iterazione del metodo di Frank Wolfe partendo dal punto (1 1), con parametri di Armijo pari a $\Delta=1$, $\gamma=\delta=\frac{1}{4}$.

Esercizio 3. Sia dato il seguente problema non vincolato:

$$\min \frac{1}{2}\left(6x_1^2+6x_1x_3+4x_3^2\right)\left(x_2^2+1\right)+(x_2-2)^2-3x_1-x_2+2x_3$$

Effettuare la migliore decomposizione in blocchi possibile e una iterazione a partire dal punto $(-1\ \ -1\ \ -1)$ di un metodo di decomposizione scelto in base alle proprietà della funzione giustificando la scelta sia della decomposizione che dell'algoritmo. Si illustrino le proprietà di convergenza dell'algoritmo scelto in questo caso.

Esercizio 4. Sia dato il seguente insieme di dati:

$$\mathcal{T}= \left\{\left(\begin{pmatrix}-1\\2\end{pmatrix},-1\right),\left(\begin{pmatrix}2\\4\end{pmatrix},-1\right),\left(\begin{pmatrix}-3\\2\end{pmatrix},-1\right),\left(\begin{pmatrix}5\\2\end{pmatrix},-1\right),\left(\begin{pmatrix}5\\4\end{pmatrix},-1\right),\right.$$
$$\left.\left(\begin{pmatrix}3\\3\end{pmatrix},1\right),\left(\begin{pmatrix}2\\0\end{pmatrix},1\right),\left(\begin{pmatrix}1\\-2\end{pmatrix},1\right),\left(\begin{pmatrix}5\\-1\end{pmatrix},1\right),\left(\begin{pmatrix}6\\1\end{pmatrix},1\right)\right\}$$

e si supponga di voler costruire una SVM per risolvere il corrispondente problema di classicazione utilizzando il kernel lineare ($k(x_i,x_j)=x_i^Tx_j$).

(i) Scrivere il problema duale che deve essere risolto per addestrare la SVM (con parametro $C = 10$).

(ii) Sapendo che la soluzione del problema duale ‘e la seguente (approssimata alla seconda cifra decimale):

$$\lambda = \begin{pmatrix} 5.56 \\ 0 \\ 0 \\ 10 \\ 0 \\ 10 \\ 4.77 \\ 0 \\ 0 \\ 0.83 \end{pmatrix}$$

individuare i vettori di supporto, calcolare l'iperpiano di separazione ottimo $w^T x + b$ e dire quali valori delle variabili ξ_i possono essere determinati.

(iii) Come classifica la SVM così addestrata il punto $x^T = (-3\ 8)$?

(iv) E' possibile dedurre la separabilità lineare dei due insiemi di dati?Come si calcola l'errore sul training set?

(iv) Addestrando nuovamente la SVM con kernel polinomiale con $\gamma = 0$ e $p = 3$, la soluzione diventa

$$\lambda = \begin{pmatrix} 0.003 \\ 0.006 \\ 0 \\ 0.009 \\ 0 \\ 0.014 \\ 0 \\ 0 \\ 0 \\ 0.004 \end{pmatrix}$$

Che cosa si può concludere guardando questa soluzione?(non è necessario eseguire nessun calcolo)

Esercizio 5 Sia dato il seguente problema:

$$\begin{array}{ll} \min & f(\alpha) \\ & \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = 0 \\ & 0 \le \alpha_i \le 3 \quad i = 1, \ldots, 4 \end{array} \tag{2}$$

Supponiamo che nel punto $\bar{\alpha} = (3\ 1\ 0\ 2)$, il gradiente della funzione valga:

$$\nabla f(\bar{\alpha}) = \begin{pmatrix} 3 \\ \beta \\ -1 \\ \gamma \end{pmatrix}$$

(i) Per quali valori di γ e β il punto $\bar{\alpha}$ è ottimo?

(ii) Per quali valori di γ e β il working set di SVM$_{light}$ è $\{2, 1\}$?