

UNIVERSITY OF PISA

Risoluzione di un problema di Constrained Convex
Quadratic Programming con Primal-Dual Interior
Point Method

Computational Mathematics for Learning
and Data Analysis project report

Cornacchia Giuliano, Salinas Mario Leonardo
Gruppo 21

INDICE

1	Introduzione	3
2	Descrizione del problema	3
3	Metodo Risolutivo	5
3.1	Primal-Dual Interior Point method	5

1 INTRODUZIONE

Nel seguente report viene descritto ed analizzato il problema numero 3 *noML* assegnato per il progetto finale di *Computational Mathematics for Learning and Data Analysis* e descritta la soluzione da noi proposta, insieme alle varie scelte implementative.

2 DESCRIZIONE DEL PROBLEMA

Date le matrici $Q \in \mathbb{R}^{n \times n}$ ed $A \in \{0,1\}^{k \times n}$, con $Q \geq 0$, e i vettori $q \in \mathbb{R}^n$ e $b = [1]^k$, il problema di ottimizzazione quadratica convessa primale P è definito come:

$$P := \begin{cases} \min x^\top Q x + q^\top x \\ Ax = b \\ x \geq 0 \end{cases} \quad (2.1)$$

dove i vincoli nella matrice A formano k semplici disgiunti della forma:

$$\sum_{i \in I^h} x_i = 1, h = 1, \dots, k \quad (2.2)$$

con $I^h, h = 1, \dots, k$ insiemi di indici che formano una partizione di $\{1, \dots, n\}$. Quindi il generico elemento della matrice A $a_{vi} = 1 \iff i \in I^v$.

Data la funzione da minimizzare in 2.1 $f(x) = x^\top Q x + q^\top x$ e la sua *funzione Lagrangiana* $L(x, \lambda_{eq}, \lambda_s)$, a P si può associare il seguente *Problema Duale di Wolfe D*:

$$D := \begin{cases} \max L(x, \lambda_{eq}, \lambda_s) \\ \nabla_x L(x, \lambda_{eq}, \lambda_s) = 0 \\ \lambda_s \geq 0 \end{cases} = \begin{cases} \max x^\top Q x + q^\top x + \lambda_{eq}^\top (Ax - b) + \lambda_s^\top (-x) \\ 2Qx + q + A^\top \lambda_{eq} - \lambda_s = 0 \\ \lambda_s \geq 0 \end{cases} \quad (2.3)$$

Possiamo a questo punto scrivere il sistema KKT associato a P .

$$\begin{cases} \nabla_x L(x, \lambda_{eq}, \lambda_s) = 0 \\ Ax - b = 0 \\ x_i \lambda_{s_i} = 0 \quad i = 1, \dots, n \\ (x, \lambda_s) \geq 0 \end{cases} \quad (2.4)$$

Scegliamo dunque di risolvere il sistema 2.4 applicando il metodo Primal-Dual Interior Point (PDIP): riformuliamo le condizioni di ottimalità 2.4 definendo una funzione $F: \mathbb{R}^{2n+k} \rightarrow \mathbb{R}^{2n+k}$ [1]:

$$F(x, \lambda_{eq}, \lambda_s) = \begin{bmatrix} \nabla_x L(x, \lambda_{eq}, \lambda_s) \\ Ax - b \\ XSe \end{bmatrix} = 0 \quad (2.5a)$$

$$(x, \lambda_s) \geq 0 \quad (2.5b)$$

dove

$$X = \text{diag}(x_1, \dots, x_n) \quad S = \text{diag}(\lambda_{s_1}, \dots, \lambda_{s_n}) \quad e^\top = [1, \dots, 1] \in \mathbb{R}^n \quad (2.6)$$

Il metodo PDIP, ad ogni iterazione k , genera triple $(x^k, \lambda_{eq}^k, \lambda_s^k)$ che soddisfano *strettamente* la 2.5b. La procedura con la quale si ricercano le direzioni $(\Delta x, \Delta \lambda_{eq}, \Delta \lambda_s)$ prende origine dal metodo di Newton per equazioni non lineari [1]: alla k -esima iterazione il metodo di Newton forma un modello lineare di F attorno al punto corrente $(x^k, \lambda_{eq}^k, \lambda_s^k)$ e ottiene le direzioni di ricerca risolvendo il seguente sistema lineare:

$$J(x, \lambda_{eq}, \lambda_s) \begin{bmatrix} \Delta x \\ \Delta \lambda_{eq} \\ \Delta \lambda_s \end{bmatrix} = -F(x, \lambda_{eq}, \lambda_s) \quad (2.7)$$

dove J è la Jacobiana di F . Nel nostro caso il sistema da risolvere diventa:

$$\begin{bmatrix} 2Q & A^\top & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda_{eq} \\ \Delta \lambda_s \end{bmatrix} = - \begin{bmatrix} r_d \\ r_p \\ XSe \end{bmatrix} \quad (2.8)$$

dove

$$r_d = 2Qx + qA^\top \lambda_{eq} - \lambda_s \quad r_p = Ax - b \quad (2.9)$$

Percorrere un passo intero lungo le direzioni trovate risolvendo 2.8 potrebbe violare il vincolo $(x, \lambda_s) \geq 0$, quindi aggiungiamo una parametro $\alpha \in (0, 1]$, detto *step-size*, che servirà a ridurre l'ampiezza del passo $(\Delta x, \Delta \lambda_{eq}, \Delta \lambda_s)$ per garantire il soddisfacimento del vincolo 2.5b.

$$(x^{k+1}, \lambda_{eq}^{k+1}, \lambda_s^{k+1}) = (x^k, \lambda_{eq}^k, \lambda_s^k) + \alpha(\Delta x, \Delta \lambda_{eq}, \Delta \lambda_s) \quad (2.10)$$

Data la corrente iterazione $(x^k, \lambda_{eq}^k, \lambda_s^k)$, che soddisfa 2.5b, introduciamo il *centering parameter* $\sigma \in [0, 1]$ e la *duality measure* $\mu = \frac{x^\top \lambda_s}{n}$; questi due parametri vengono utilizzati per direzionare il Newton step verso un punto per il quale valga $x_i \lambda_{s_i} = \sigma \mu$, piuttosto che a una soluzione diretta di 2.4. A seguito di questa considerazione, il nuovo step verrà calcolato risolvendo il *KKT perturbato* definito come:

$$\begin{bmatrix} 2Q & A^\top & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda_{eq} \\ \Delta \lambda_s \end{bmatrix} = - \begin{bmatrix} r_d \\ r_p \\ XSe - \sigma \mu e \end{bmatrix} \quad (2.11)$$

Il sistema 2.11 è non-simmetrico, lo trasformiamo in un sistema lineare equivalente simmetrico, eliminando la terza riga ed esprimendo $\Delta \lambda_s$ in funzione di Δx .

La terza riga di 2.11 può essere eliminata poichè durante le iterazioni x_i ed s_i rimangono strettamente positivi; per ricavare $\Delta \lambda_s$ in funzione di Δx dobbiamo prima isolare $\Delta \lambda_s$ sempre dalla terza riga del sistema:

$$\begin{aligned} S\Delta x + X\Delta \lambda_s &= -XSe + \sigma \mu e \\ X\Delta \lambda_s &= \sigma \mu e - XSe - S\Delta x \\ \Delta \lambda_s &= X^{-1}(\sigma \mu e - XSe - S\Delta x) - \lambda_s \end{aligned} \quad (2.12)$$

possiamo dunque riscrivere 2.11 sostituendo $\Delta \lambda_s$ come in 2.12:

$$\begin{aligned} 2Q\Delta x + A^\top \Delta \lambda_{eq} - \Delta \lambda_s &= -2Qx - q - A^\top \lambda_{eq} + \lambda_s \\ 2Q\Delta x + A^\top \Delta \lambda_{eq} - X^{-1}(\sigma \mu e - XSe - S\Delta x) - \cancel{\lambda_s} &= -2Qx - q - A^\top \lambda_{eq} + \cancel{\lambda_s} \\ (2Q + X^{-1}S)\Delta x + A^\top \Delta \lambda_{eq} &= -2Qx - q - A^\top \lambda_{eq} + X^{-1}\sigma \mu e \end{aligned} \quad (2.13)$$

ponendo $M = 2Q + X^{-1}S$ otteniamo il seguente sistema simmetrico detto anche *augmented KKT*:

$$\begin{bmatrix} 2Q + X^{-1}S & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda_{eq} \end{bmatrix} = - \begin{bmatrix} r_d - X^{-1}\sigma\mu e \\ r_p \end{bmatrix} \quad (2.14a)$$

$$\Delta \lambda_s = X^{-1}(\sigma\mu e - S\Delta x) - \lambda_s \quad (2.14b)$$

la matrice a sinistra in 2.14a è simmetrica poichè:

- M è simmetrica perchè somma di una matrice simmetrica e una matrice diagonale
- il blocco inferiore sinistro e superiore destro sono l'uno il trasposto dell'altro

inoltre se A ha rango massimo essa è non-singolare, e 2.14a ammette soluzione. I vincoli in A , nel nostro caso di studio, formano k semplici disgiunti, quindi $rank(A) = k$.

Il sistema 2.14a è dunque simmetrico, sparso e la matrice potrebbe essere mal condizionata a causa del prodotto $X^{-1}S$. Il metodo solitamente utilizzato per risolvere sistemi simmetrici sparsi è MINRES, ma non essendo stato affrontato durante il corso, utilizzeremo una sua generalizzazione, GMRES.

Avremmo potuto anche optare per LU-factorization, ma avrebbe sofferto di gravi problemi di stabilità sulle nostre istanze.

3 METODO RISOLUTIVO

3.1 PRIMAL-DUAL INTERIOR POINT METHOD

L'intuizione principale dei metodi Primal-Dual è quella di considerare sia il problema di minimizzazione P che il suo duale D per ottenere un limite superiore ed inferiore della soluzione. Da P e D si ricava quindi il sistema KKT associato 2.14a

RIFERIMENTI BIBLIOGRAFICI

- [1] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2 edition, 2006.