Metodi iterativi di Bregman per la minimizzazione ℓ_1

Alessandro La Farciola

Università di Pisa

13 Settembre 2021

Referenze

- S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
- W. Yin, S. Osher, D. Goldfarb, J. Darbon, *Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing*, SIAM J. Imaging Sci., 1 (2008), pp. 143–168.
- T. Goldstein, S. Osher, *The Split Bregman Method for L1-Regularized Problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.

Formulazione del problema

Molti importanti problemi di ingegneria, computer science o imaging science possono essere ricondotti a problemi di ottimizzazione del tipo

$$\min_{u \in \mathbb{R}^n} \{ \|u\|_1 \ : \ Au = f \} \qquad (\textit{Basis Pursuit})$$

dove $A \in \mathbb{R}^{m \times n}$ e $f \in \mathbb{R}^m$.

Il sistema $Au = f$ è tipicamente sottodeterminato, cioè presi $m < n$, può avere più di una soluzione.

Per la risoluzione di problemi di questo tipo, come la ricostruzione di segnali, si parla di *Compressed sensing* (CS).

Formulazione del problema

Molti importanti problemi di ingegneria, computer science o imaging science possono essere ricondotti a problemi di ottimizzazione del tipo

$$\min_{u \in \mathbb{R}^n} \{\|u\|_1 \ : \ Au = f\} \qquad (\textit{Basis Pursuit})$$

dove $A \in \mathbb{R}^{m \times n}$ e $f \in \mathbb{R}^m$.

Il sistema $Au = f$ è tipicamente sottodeterminato, cioè presi $m < n$, può avere più di una soluzione.

Per la risoluzione di problemi di questo tipo, come la ricostruzione di segnali, si parla di *Compressed sensing* (CS).

Università di Pisa

Formulazione del problema

Molti importanti problemi di ingegneria, computer science o imaging science possono essere ricondotti a problemi di ottimizzazione del tipo

$$\min_{u \in \mathbb{R}^n} \{\|u\|_1 \ : \ Au = f\} \qquad (\textit{Basis Pursuit})$$

dove $A \in \mathbb{R}^{m \times n}$ e $f \in \mathbb{R}^m$.

Il sistema $Au = f$ è tipicamente sottodeterminato, cioè presi $m < n$, può avere più di una soluzione.

Per la risoluzione di problemi di questo tipo, come la ricostruzione di segnali, si parla di *Compressed sensing* (CS).

Formulazione del problema

CS è basata sull'idea di ricostruire il segnale da una piccola quantità di dati. Dato $\bar{u} \in \mathbb{R}^n$ un segnale molto sparso, ovvero

$$k = \|\bar{u}\|_0 = |\{i : \bar{u}_i \neq 0\}| \ll n,$$

allora è possibile considerare trasformazione lineare $f = A\bar{u} \in \mathbb{R}^m$ per $k < m \ll n$, e ricavare $\bar{u}$ da f risolvendo il problema iniziale.

Formulazione del problema

CS è basata sull'idea di ricostruire il segnale da una piccola quantità di dati. Dato $\bar{u} \in \mathbb{R}^n$ un segnale molto sparso, ovvero

$$k = \|\bar{u}\|_0 = |\{i : \bar{u}_i \neq 0\}| \ll n,$$

allora è possibile considerare trasformazione lineare $f = A\bar{u} \in \mathbb{R}^m$ per $k < m \ll n$, e ricavare $\bar{u}$ da f risolvendo il problema iniziale.

Infine, invece di risolvere il problema vincolato di prima, è preferibile risolvere il seguente problema di ottimizzazione composta non vincolato

$$\min_{u \in \mathbb{R}^n} \mu \|u\|_1 + \frac{1}{2} \|Au - f\|_2^2$$

con $\mu \in \mathbb{R}$. Una delle possibili tecniche di risoluzione è ISTA.

UNIVERSITÀ DI PISA

Regolarizzazione mediante le iterate di Bregman

Consideriamo il seguente problema classico di recupero immagini: data un immagine con rumore $f : \Omega \longrightarrow \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^2$ aperto limitato, si vuole ottenere una decomposizione del tipo

$$f = u + v$$

dove u è il segnale vero e v il rumore.

Regolarizzazione mediante le iterate di Bregman

Consideriamo il seguente problema classico di recupero immagini: data un immagine con rumore $f : \Omega \longrightarrow \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^2$ aperto limitato, si vuole ottenere una decomposizione del tipo

$$f = u + v$$

dove u è il segnale vero e v il rumore.

Uno dei metodi più diffusi per approssimare la soluzione si deve a Rudin, Osher, Fatemi (ROF) e si ottiene risolvendo

$$\min_{u \in BV(\Omega)} \left\{ \int_\Omega |\nabla u| + \lambda \, \|f - u\|_{L^2}^2 \right\}$$

per qualche $\lambda > 0$ e con $BV(\Omega)$, lo spazio delle funzioni a variazione limitata su Ω.

Distanza di Bregman

Definizione

Si chiama *distanza di Bregman* rispetto ad una funzione convessa $J(\cdot)$ tra due punti la quantità

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v \rangle$$

dove $p \in \partial J(v)$.

Università di Pisa

Distanza di Bregman

Definizione

Si chiama *distanza di Bregman* rispetto ad una funzione convessa $J(\cdot)$ tra due punti la quantità

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v \rangle$$

dove $p \in \partial J(v)$.

La quantità $D_J^p(u, v)$ **non** è una distanza nel significato usuale, infatti $D_J^p(u, v) \neq D_J^p(v, u)$ in generale e non è sempre verificata la disuguaglianza triangolare. In ogni caso, misura la *vicinanza* tra u e v nel senso che $D_J^p(u, v) \geq 0$, e $D_J^p(u, v) \geq D_J^p(w, v)$ per ogni w nel segmento che congiunge u e v.

Regolarizzazione mediante le iterate di Bregman

Poniamo $J(u) = \int_\Omega |\nabla u|$, la *regolarizzazione mediante le iterate di Bregman* consiste nel risolvere una serie di problemi di ottimizzazione del tipo

$$
\begin{aligned}
u^{k+1} &= \underset{u \in BV(\Omega)}{\arg\min} \left\{ D_J^{p^k}(u, u^k) + \lambda \|f - u\|_{L^2}^2 \right\} \\
&= \underset{u \in BV(\Omega)}{\arg\min} \left\{ J(u) - \langle p^k, u - u^k \rangle + \lambda \|f - u\|_{L^2}^2 \right\}
\end{aligned}
$$

con $k = 0, 1, \ldots$, partendo da $u^0 = 0$ e $p^0 = 0$.

Università di Pisa

Regolarizzazione mediante le iterate di Bregman

Poniamo $J(u) = \int_\Omega |\nabla u|$, la *regolarizzazione mediante le iterate di Bregman* consiste nel risolvere una serie di problemi di ottimizzazione del tipo

$$\begin{aligned} u^{k+1} &= \underset{u \in BV(\Omega)}{\arg\min} \left\{ D_J^{p^k}(u, u^k) + \lambda \|f - u\|_{L^2}^2 \right\} \\ &= \underset{u \in BV(\Omega)}{\arg\min} \left\{ J(u) - \langle p^k, u - u^k \rangle + \lambda \|f - u\|_{L^2}^2 \right\} \end{aligned}$$

con $k = 0, 1, \ldots$, partendo da $u^0 = 0$ e $p^0 = 0$.

- Per $k = 0$ stiamo risolvendo il problema originale.
- Dall'ottimalità di u^{k+1} segue che $0 \in \partial J(u^{k+1}) - p^k + u^{k+1} - f$ e quindi possiamo porre

$$p^{k+1} := p^k + f - u^{k+1} \quad \in \partial J(u^{k+1})$$

e iterare il procedimento.

Università di Pisa

Regolarizzazione mediante le iterate di Bregman

Poniamo $J(u) = \int_\Omega |\nabla u|$, la *regolarizzazione mediante le iterate di Bregman* consiste nel risolvere una serie di problemi di ottimizzazione del tipo

$$\begin{aligned} u^{k+1} &= \operatorname*{arg\,min}_{u \in BV(\Omega)} \left\{ D_J^{p^k}(u, u^k) + \lambda \|f - u\|_{L^2}^2 \right\} \\ &= \operatorname*{arg\,min}_{u \in BV(\Omega)} \left\{ J(u) - \langle p^k, u - u^k \rangle + \lambda \|f - u\|_{L^2}^2 \right\} \end{aligned}$$

con $k = 0, 1, \ldots$, partendo da $u^0 = 0$ e $p^0 = 0$.

- Per $k = 0$ stiamo risolvendo il problema originale.
- Dall'ottimalità di u^{k+1} segue che $0 \in \partial J(u^{k+1}) - p^k + u^{k+1} - f$ e quindi possiamo porre

$$p^{k+1} := p^k + f - u^{k+1} \quad \in \partial J(u^{k+1})$$

  e iterare il procedimento.

Risultati di convergenza

Consideriamo il seguente problema generale

$$\min_u J(u) + H(u)$$

con $J(\cdot)$ una funzione convessa e $H(\cdot)$ una funzione convessa e differenziabile.
Allora, l' algoritmo di *regolarizzazione mediante le iterate di Bregman* diventa

$$u^{k+1} = \arg\min_u \left\{ D_J^{p^k}(u, u^k) + H(u) \right\}$$
$$p^{k+1} = p^k - \nabla H(u^{k+1}) \in \partial J(u^{k+1})$$

UNIVERSITÀ DI PISA

Risultati di convergenza

Teorema

Siano $J(\cdot)$ una funzione convessa e $H(\cdot)$ una funzione convessa e differenziabile. Supponiamo che le soluzioni u^{k+1} dei problemi di ottimizzazione alla k-esima iterazione esistono. Allora, la successione $\{u^k\}$ soddisfa

1. *Monotonia discendente in H:* $H(u^{k+1}) \leq H(u^{k+1}) + D_J^{p^k}(u^{k+1}, u^k) \leq H(u^k)$.
2. *Se $\tilde{u}$ minimizza $H(\cdot)$ e $J(\tilde{u}) < \infty$, allora $H(u^k) \leq H(\tilde{u}) + J(\tilde{u})/k$.*
3. *Sia $H(\cdot) = H(\cdot, f)$ e supponiamo che $H(\tilde{u}, f) \leq \delta^2$ e $H(\tilde{u}, g) = 0$, dove $f, g, \tilde{u}, \delta$ rappresentano i dati con rumore, i dati senza rumore, la soluzione esatta, il livello di rumore, rispettivamente. Allora*

$$D_J^{p^{k+1}}(\tilde{u}, u^{k+1}) < D_J^{p^k}(\tilde{u}, u^k)$$

finché $H(u^{k+1}, f) > \delta^2$.

Teorema

Siano $J(\cdot)$ una funzione convessa e $H(\cdot)$ una funzione convessa e differenziabile. Supponiamo che le soluzioni u^{k+1} dei problemi di ottimizzazione alla k-esima iterazione esistono. Allora, la successione $\{u^k\}$ soddisfa

1. *Monotonia discendente in H:* $H(u^{k+1}) \leq H(u^{k+1}) + D_J^{p^k}(u^{k+1}, u^k) \leq H(u^k)$.
2. *Se $\tilde{u}$ minimizza $H(\cdot)$ e $J(\tilde{u}) < \infty$, allora $H(u^k) \leq H(\tilde{u}) + J(\tilde{u})/k$.*
3. *Sia $H(\cdot) = H(\cdot, f)$ e supponiamo che $H(\tilde{u}, f) \leq \delta^2$ e $H(\tilde{u}, g) = 0$, dove $f, g, \tilde{u}, \delta$ rappresentano i dati con rumore, i dati senza rumore, la soluzione esatta, il livello di rumore, rispettivamente. Allora*

$$D_J^{p^{k+1}}(\tilde{u}, u^{k+1}) < D_J^{p^k}(\tilde{u}, u^k)$$

finché $H(u^{k+1}, f) > \delta^2$.

Risultati di convergenza

Teorema

Siano $J(\cdot)$ *una funzione convessa e* $H(\cdot)$ *una funzione convessa e differenziabile. Supponiamo che le soluzioni* u^{k+1} *dei problemi di ottimizzazione alla k-esima iterazione esistono. Allora, la successione* $\{u^k\}$ *soddisfa*

1. *Monotonia discendente in H:* $H(u^{k+1}) \leq H(u^{k+1}) + D_J^{p^k}(u^{k+1}, u^k) \leq H(u^k)$.
2. *Se* $\tilde{u}$ *minimizza* $H(\cdot)$ *e* $J(\tilde{u}) < \infty$, *allora* $H(u^k) \leq H(\tilde{u}) + J(\tilde{u})/k$.
3. *Sia* $H(\cdot) = H(\cdot, f)$ *e supponiamo che* $H(\tilde{u}, f) \leq \delta^2$ *e* $H(\tilde{u}, g) = 0$, *dove* $f, g, \tilde{u}, \delta$ *rappresentano i dati con rumore, i dati senza rumore, la soluzione esatta, il livello di rumore, rispettivamente. Allora*

$$D_J^{p^{k+1}}(\tilde{u}, u^{k+1}) < D_J^{p^k}(\tilde{u}, u^k)$$

finché $H(u^{k+1}, f) > \delta^2$.

Applicazione delle iterate di Bregman

Ricordiamo il problema iniziale

$$\min_{u \in \mathbb{R}^n} \{ \|u\|_1 \ : \ Au = f \} \qquad (\textit{Basis Pursuit})$$

dove $A \in \mathbb{R}^{m \times n}$ e $f \in \mathbb{R}^m$. O equivalentemente,

$$\min_{u \in \mathbb{R}^n} \mu \|u\|_1 + \frac{1}{2} \|Au - f\|_2^2$$

Si vuole, ora, applicare il procedimento mediante le iterate di Bregman appena visto per risoluzione di questo problema.

Applicazione delle iterate di Bregman

Ricordiamo il problema iniziale

$$\min_{u\in\mathbb{R}^n}\{\|u\|_1 \;:\; Au = f\} \qquad (\textit{Basis Pursuit})$$

dove $A \in \mathbb{R}^{m\times n}$ e $f \in \mathbb{R}^m$. O equivalentemente,

$$\min_{u\in\mathbb{R}^n} \mu \|u\|_1 + \frac{1}{2}\|Au - f\|_2^2$$

Si vuole, ora, applicare il procedimento mediante le iterate di Bregman appena visto per risoluzione di questo problema.

Università di Pisa

Fomulazione algoritmi

L'algoritmo generale era

$$u^{k+1} = \arg\min_u \left\{ D_J^{p^k}(u, u^k) + H(u) \right\}$$
$$p^{k+1} = p^k - \nabla H(u^{k+1}) \in \partial J(u^{k+1})$$

Prendendo $J(u) = \mu \|u\|_1$ and $H(u) = \frac{1}{2}\|Au - f\|_2^2$, l'algoritmo diventa **Algoritmo 1**:

$$u^0 = 0, \;\; p^0 = 0$$
$$\textit{For } k = 0, 1, \ldots \;\; \textit{do}$$
$$u^{k+1} = \arg\min_u D_J^{p^k}(u, u^k) + \frac{1}{2}\|Au - f\|_2^2$$
$$p^{k+1} = p^k - A^T(Au^{k+1} - f) \;\; \in \partial J(u^{k+1})$$

dato che $\nabla H(u^{k+1}) = A^T(Au^{k+1} - f)$.

UNIVERSITÀ DI PISA

Fomulazione algoritmi

L'algoritmo generale era

$$u^{k+1} = \arg\min_u \left\{ D_J^{p^k}(u, u^k) + H(u) \right\}$$
$$p^{k+1} = p^k - \nabla H(u^{k+1}) \in \partial J(u^{k+1})$$

Prendendo $J(u) = \mu \|u\|_1$ and $H(u) = \frac{1}{2}\|Au - f\|_2^2$, l'algoritmo diventa
Algoritmo 1:

$$u^0 = 0, \;\; p^0 = 0$$
$$\textit{For} \;\; k = 0, 1, \ldots \;\; \textit{do}$$
$$u^{k+1} = \arg\min_u D_J^{p^k}(u, u^k) + \frac{1}{2}\|Au - f\|_2^2$$
$$p^{k+1} = p^k - A^T(Au^{k+1} - f) \;\; \in \partial J(u^{k+1})$$

dato che $\nabla H(u^{k+1}) = A^T(Au^{k+1} - f)$.

Università di Pisa

Fomulazione algoritmi

Ora, invece, consideriamo

Algoritmo 2:

$$
\begin{aligned}
& u^0 = 0, \;\; f^0 = 0 \\
& \textit{For } k = 0, 1, \dots \;\; \textit{do} \\
& f^{k+1} = f + (f^k - Au^k) \\
& u^{k+1} = \arg\min_u J(u) + \frac{1}{2} \left\| Au - f^{k+1} \right\|_2^2
\end{aligned}
$$

Università di Pisa

Ora, invece, consideriamo

Algoritmo 2:

$$u^0 = 0,\ \ f^0 = 0$$
$$\textit{For}\ \ k = 0, 1, \ldots\ \ \textit{do}$$
$$f^{k+1} = f + (f^k - Au^k)$$
$$u^{k+1} = \arg\min_u J(u) + \frac{1}{2}\left\|Au - f^{k+1}\right\|_2^2$$

E' possibile notare che i due Algoritmi sono equivalenti. Infatti, vale il seguente Teorema.

Teorema

L' Algoritmo 1 e l' Algoritmo 2 sono equivalenti nel senso che alla k-esima iterazione la funzione obiettivo è la medesima (a meno di costante).

Risultati di convergenza

Teorema

Supponiamo che un certo u^k soddisfa $Au^k = f$; allora u^k è soluzione del problema

$$\min_{u \in \mathbb{R}^n} \{ \|u\|_1 \; : \; Au = f \} \qquad (\textit{Basis Pursuit}).$$

Università di Pisa

Risultati di convergenza

Teorema

Supponiamo che un certo u^k soddisfa $Au^k = f$; allora u^k è soluzione del problema

$$\min_{u \in \mathbb{R}^n} \{ \|u\|_1 \ : \ Au = f \} \qquad (\textit{Basis Pursuit}).$$

Teorema

Esiste un finito K tale che ciascun u^k con $k \geq K$ soddisfa $Au^k = f$.

Università di Pisa

Equivalenza con il metodo della Lagrangiana aumentata

Consideriamo il problema di ottimizzazione

$$\min_u s(u) \quad \textit{vincolata a } c_i(u) = 0, \quad i = 1, \ldots, m,$$

il metodo della Lagrangiana aumentata consiste nel minimizzare

$$L(u, \lambda^k, \nu) := s(u) + \sum_{i=1}^{m} \lambda_i^k c_i(u) + \frac{1}{2} \sum_{i=1}^{m} \nu_i c_i^2(u)$$

rispetto a u ad ogni iterazione k e, chiamato questo u^{k+1}, aggiornare i moltiplicatori in modo che

$$\lambda_i^{k+1} = \lambda_i^k + \nu_i c_i(u^{k+1}).$$

Università di Pisa

Possiamo, ora, vedere che questo metodo è equivalente all' *Algoritmo 1* di prima prendendo

$$s(u) = J(u)$$

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = Au - f$$

$$p^k = -A^T \lambda^k$$

$$\nu_i \equiv 1 \quad \forall i.$$

Equivalenza con il metodo della Lagrangiana aumentata

A questo punto, valgono i seguenti passaggi che garantiscono l'equivalenza prendendo $u^0 = 0$ e $\lambda^0 = 0$.

$$
\begin{aligned}
L(u, \lambda^k, \nu) &= s(u) + \sum_{i=1}^{m} \lambda_i^k c_i(u) + \frac{1}{2} \sum_{i=1}^{m} \nu_i c_i^2(u) \\
&= J(u) + \langle \lambda^k, Au \rangle + \frac{1}{2} \|Au - f\|_2^2 + C_1 \\
&= J(u) - \langle p^k, u \rangle + \frac{1}{2} \|Au - f\|_2^2 + C_1 \\
&= D_J^{p^k}(u, u^k) + \frac{1}{2} \|Au - f\|_2^2 + C_2.
\end{aligned}
$$

Università di Pisa

Consideriamo il problema

$$\min_u E(u) \textit{ tale che } Au = f,$$

dove $E(\cdot)$ è una funzione convessa, A un operatore lineare e $f \in \mathbb{R}^m$.

Tale problema è equivalente a risolvere il seguente problema di ottimizzazione non vincolato:

$$\min_u E(u) + \frac{\lambda}{2} \|Au - f\|_2^2 .$$

UNIVERSITÀ DI PISA

Ottimizzazione vincolata via iterate di Bregman

Consideriamo il problema

$$\min_u E(u) \textit{ tale che } Au = f,$$

dove $E(\cdot)$ è una funzione convessa, A un operatore lineare e $f \in \mathbb{R}^m$.

Tale problema è equivalente a risolvere il seguente problema di ottimizzazione non vincolato:

$$\min_u E(u) + \frac{\lambda}{2} \|Au - f\|_2^2 .$$

Ottimizzazione vincolata via iterate di Bregman

Applichiamo, ora, il metodo delle iterate di Bregman (Algoritmo 1) e otteniamo:

$$
\begin{aligned}
u^{k+1} &= \arg\min_u D_E^{p^k}(u, u^k) + \frac{\lambda}{2} \|Au - f\|_2^2 \\
p^{k+1} &= p^k - \lambda A^T (Au^{k+1} - f).
\end{aligned}
$$

UNIVERSITÀ DI PISA

Applichiamo, ora, il metodo delle iterate di Bregman (Algoritmo 1) e otteniamo:

$$\begin{aligned} u^{k+1} &= \arg\min_u D_E^{p^k}(u, u^k) + \frac{\lambda}{2}\|Au - f\|_2^2 \\ p^{k+1} &= p^k - \lambda A^T(Au^{k+1} - f). \end{aligned}$$

Se, invece, applichiamo l'Algoritmo 2, che sappiamo essere equivalente, si ha che

$$\begin{aligned} f^{k+1} &= f + f^k - Au^k \\ u^{k+1} &= \arg\min_u E(u) + \frac{\lambda}{2}\left\|Au - f^{k+1}\right\|_2^2 \end{aligned}$$

Risultato di convergenza

A differenza del risultato di convergenza dell'articolo precedente dove si era assunta uno specifica funzione obiettivo, qui lo stesso risultato viene esteso.

Teorema

Sia $E : \mathbb{R}^n \longrightarrow \mathbb{R}$ una funzione convessa, sia $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ un operatore lineare e consideriamo l' Algoritmo 2 precedente. Supponiamo che una certa iterata, $u^\star$, soddisfi $Au^\star = f$. Allora, $u^\star$ è una soluzione del problema vincolato iniziale:

$$\min_u E(u) + \frac{\lambda}{2} \|Au - f\|_2^2 .$$

Università di Pisa

Risultato di convergenza

A differenza del risultato di convergenza dell'articolo precedente dove si era assunta uno specifica funzione obiettivo, qui lo stesso risultato viene esteso.

Teorema

Sia $E : \mathbb{R}^n \longrightarrow \mathbb{R}$ *una funzione convessa, sia* $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *un operatore lineare e consideriamo l' Algoritmo 2 precedente. Supponiamo che una certa iterata,* $u^\star$*, soddisfi* $Au^\star = f$*. Allora,* $u^\star$ *è una soluzione del problema vincolato iniziale:*

$$\min_u E(u) + \frac{\lambda}{2} \|Au - f\|_2^2 .$$

Università di Pisa

Il metodo *Split Bregman*

Consideriamo il problema generale di ottimizzazione ℓ_1

$$\min_{u} \|\Phi(u)\|_{L^1} + H(u)$$

dove sia $\Phi(\cdot)$ sia $H(\cdot)$ sono funzioni convesse. Assumiamo inoltre che $\Phi(\cdot)$ sia differenziabile.

Ora, invece di considerare il precedente problema, prendiamo l'equivalente

$$\min_{u,d} \|d\|_1 + H(u) \quad \textit{tale che } d = \Phi(u)$$

Da qui, come prima, possiamo ricavare il problema non vincolato seguente:

$$\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \|d - \Phi(u)\|_2^2 .$$

Università di Pisa

Il metodo *Split Bregman*

Consideriamo il problema generale di ottimizzazione ℓ_1

$$\min_u \|\Phi(u)\|_{L^1} + H(u)$$

dove sia $\Phi(\cdot)$ sia $H(\cdot)$ sono funzioni convesse. Assumiamo inoltre che $\Phi(\cdot)$ sia differenziabile.

Ora, invece di considerare il precedente problema, prendiamo l'equivalente

$$\min_{u,d} \|d\|_1 + H(u) \quad \textit{tale che } d = \Phi(u)$$

Da qui, come prima, possiamo ricavare il problema non vincolato seguente:

$$\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \|d - \Phi(u)\|_2^2 .$$

UNIVERSITÀ DI PISA

Il metodo *Split Bregman*

Consideriamo il problema generale di ottimizzazione ℓ_1

$$\min_u \|\Phi(u)\|_{L^1} + H(u)$$

dove sia $\Phi(\cdot)$ sia $H(\cdot)$ sono funzioni convesse. Assumiamo inoltre che $\Phi(\cdot)$ sia differenziabile.

Ora, invece di considerare il precedente problema, prendiamo l'equivalente

$$\min_{u,d} \|d\|_1 + H(u) \quad \textit{tale che } d = \Phi(u)$$

Da qui, come prima, possiamo ricavare il problema non vincolato seguente:

$$\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \|d - \Phi(u)\|_2^2 .$$

Il metodo *Split Bregman*

Poniamo $E(u,d) = \|d\|_1 + H(u)$ e $A(u,d) := d - \Phi(u)$; allora, il problema diventa un'applicazione di quanto detto in precedenza, ovvero

$$\min_u E(u) \text{ tale che } Au = f,$$

con $f = 0$.

Il metodo *Split Bregman*

Poniamo $E(u,d) = \|d\|_1 + H(u)$ e $A(u,d) := d - \Phi(u)$; allora, il problema diventa un'applicazione di quanto detto in precedenza, ovvero

$$\min_u E(u) \textit{ tale che } Au = f,$$

con $f = 0$.

A questo punto, applicando il metodo delle iterate di Bregman (Algoritmo 1) otteniamo

$$\begin{aligned}
(u^{k+1}, d^{k+1}) &= \arg\min_{u,d} D_E^{p^k}(u, u^k, d, d^k) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \\
&= \arg\min_{u,d} E(u,d) - \langle p_u^k, u - u^k\rangle - \langle p_d^k, d - d^k\rangle + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \\
p_u^{k+1} &= p_u^k - \lambda(\nabla\Phi)^T(\Phi(u^{k+1}) - d^{k+1}) \\
p_d^{k+1} &= p_d^k - \lambda(d^{k+1} - \Phi(u^{k+1})).
\end{aligned}$$

Se, invece, utilizziamo la semplificazione dell' Algoritmo 2 otteniamo il seguente metodo.

Split Bregman Iteration

$$f^{k+1} = f^k + (\Phi(u^k) - d^k).$$

$$(u^{k+1}, d^{k+1}) = \arg\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \left\| d - \Phi(u) - f^{k+1} \right\|_2^2.$$

Se, invece, utilizziamo la semplificazione dell' Algoritmo 2 otteniamo il seguente metodo.

Split Bregman Iteration

$$f^{k+1} = f^k + (\Phi(u^k) - d^k).$$

$$(u^{k+1}, d^{k+1}) = \arg\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \left\| d - \Phi(u) - f^{k+1} \right\|_2^2.$$

All'iterazione k-esima è possibile minimizzare la funzione obiettivo prima rispetto alla variabile u, e poi rispetto a d e quindi si ottengono:

$$\textit{Step 1}: \;\; u^{k+1} = \arg\min_{u} H(u) + \frac{\lambda}{2} \left\| d^k - \Phi(u) - f^{k+1} \right\|_2^2$$

$$\textit{Step 2}: \;\; d^{k+1} = \arg\min_{d} \|d\|_1 + \frac{\lambda}{2} \left\| d - \Phi(u^{k+1}) - f^{k+1} \right\|_2^2.$$

Il metodo *Split Bregman*

Infine, possiamo vedere che ogni punto fisso dell'algoritmo *Split Bregman Iteration* risolve il problema originario

$$\min_{u,d} \|d\|_1 + H(u) \quad \textit{tale che } d = \Phi(u).$$

Infatti se $(u^\star, f^\star)$ è un punto fisso, allora soddisfa

$$f^\star = f^\star + \Phi(u^\star) - d^\star$$

che implica

$$d^\star = \Phi(u^\star).$$

Pertanto,

$$A(u^\star, d^\star) = d^\star - \Phi(u^\star) = 0 = f$$

e, grazie all'ultimo teorema di convergenza visto, $(u^\star, f^\star)$ è una soluzione.

Università di Pisa

GRAZIE PER L'ATTENZIONE