Soluciones Esparsas de un Sistema Lineal
Algoritmos Re-Weighted ℓ_1

Matías Valdés

Encuentro de Estudiantes de Matemática de Uruguay
21 de setiembre de 2019

Problema
Minimización l_1
Problema Weighted ℓ_1
Algoritmo RWl_1 CWB
Metodología RWℓ_1 propuesta
Algoritmo RWℓ_1 subgradiente
Sparse Coding

Problema a resolver

- Sistema lineal $\Phi x = b$ con infinitas soluciones
- Hallar solución con mayor cantidad de coordenadas nulas: la más "esparsa"

Problema
Minimización l_1
Problema Weighted ℓ_1
Algoritmo RWl_1 CWB
Metodología RWℓ_1 propuesta
Algoritmo RWℓ_1 subgradiente
Sparse Coding

Problema a resolver

- Sistema lineal $\Phi x = b$ con infinitas soluciones
- Hallar solución con mayor cantidad de coordenadas nulas: la más "esparsa"

Problema de optimización:

$$\begin{aligned} \operatorname{argmin} \quad & \|x\|_0 \\ \Phi x = b & \\ x \in \mathbb{R}^n & \end{aligned} \qquad (P_0)$$

- Pseudo norma ℓ_0 cuenta coordenadas no nulas

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Problema a resolver

- Sistema lineal $\Phi x = b$ con infinitas soluciones
- Hallar solución con mayor cantidad de coordenadas nulas: la más "esparsa"

Problema de optimización:

$$\begin{aligned} \operatorname{argmin} \quad & \|x\|_0 \\ \Phi x = b & \\ x \in \mathbb{R}^n & \end{aligned} \qquad (P_0)$$

- Pseudo norma ℓ_0 cuenta coordenadas no nulas
- No existe algoritmo que resuelva (P_0) en tiempo polinomial: NP-difícil

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Problema alternativo: Minimización ℓ_1

$$\underset{\substack{\Phi x = b \\ x \in \mathbb{R}^n}}{\operatorname{argmin}} \|x\|_1 = \underset{\substack{\Phi x = b \\ x \in \mathbb{R}^n}}{\operatorname{argmin}} \sum_{i=1}^{n} |x_i| \qquad (P_1)$$

- Problema convexo $\Rightarrow$ solución en tiempo polinomial

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Problema alternativo: Minimización ℓ_1

$$\begin{array}{l} \operatorname{argmin} \\ \Phi x = b \\ x \in \mathbb{R}^n \end{array} \|x\|_1 = \begin{array}{l} \operatorname{argmin} \\ \Phi x = b \\ x \in \mathbb{R}^n \end{array} \sum_{i=1}^{n} |x_i| \qquad (P_1)$$

- Problema convexo $\Rightarrow$ solución en tiempo polinomial

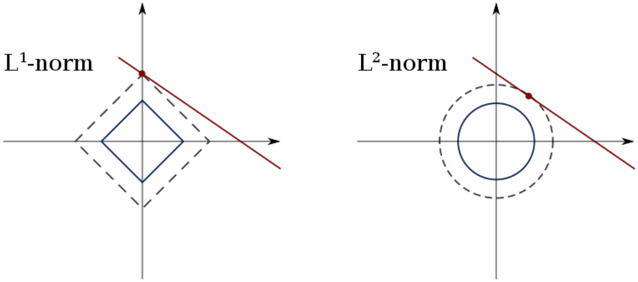


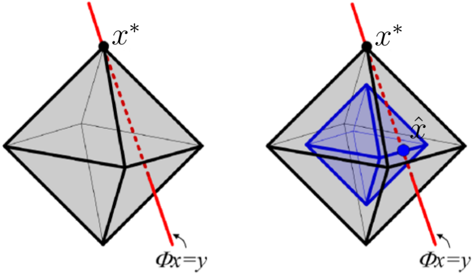
Figure: Norma ℓ_2 solución menos esparsa que con ℓ_1

Problema
Minimización l_1
Problema Weighted ℓ_1
Algoritmo RWl_1 CWB
Metodología RWℓ_1 propuesta
Algoritmo RWℓ_1 subgradiente
Sparse Coding

Ejemplo en que $P_1 \not\Rightarrow P_0$

- (P_1) no es equivalente al problema original (P_0)

$$\Phi = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \Phi x = b, \quad x \in \mathbb{R}^3.$$



(a) (b)

- $\|\hat{x}\|_1 = \|\left(\frac{1}{3}, 0, \frac{1}{3}\right)\|_1 = \frac{2}{3} < 1 = \|(0,1,0)\|_1 = \|x^*\|_1$

Problema Weighted ℓ_1 (ℓ_1 con pesos)

$$\hat{x} \in \begin{array}{c} \text{argmin} \\ \Phi x = b \\ x \in \mathbb{R}^n \end{array} \sum_{i=1}^{n} w_i |x_i| \qquad (P_1 W)$$

- Objetivo: pesos w_i tales que $\hat{x} = x^*$, solución esparsa

Problema Weighted ℓ_1 (ℓ_1 con pesos)

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

$$\hat{x} \in \underset{\substack{\Phi x = b \\ x \in \mathbb{R}^n}}{\operatorname{argmin}} \sum_{i=1}^{n} w_i |x_i| \qquad (P_1 W)$$

- Objetivo: pesos w_i tales que $\hat{x} = x^*$, solución esparsa
- Caso ideal: si se conoce solución esparsa x^*, tomar:

$$w_i = \frac{1}{|x_i^*|} \in [0, \infty]$$

- Con esto: $x_i^* = 0 \Rightarrow$ "costo" $w_i = \frac{1}{0} + \infty \Rightarrow \hat{x}_i = 0$

Ejemplo en que $P_1 \not\Rightarrow P_0$ pero $P_1 W \Rightarrow P_0$

Ejemplo anterior: $x^* = (0,1,0)$. Tomando: $w = (3,1,3)$

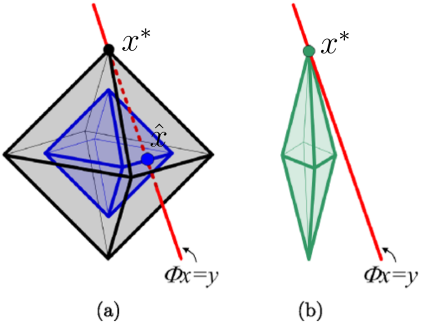


Figure: (a) x^* no es solución de (P_1); (b) x^* solución de $(P_1 W)$

Algoritmo RWℓ_1 de Candès, Wakin y Boyd [2008]

- Caso ideal $w_i = \frac{1}{|x_i^*|}$ sugiere actualizar pesos mediante:

$$w_i^{k+1} = \frac{1}{|x_i^{k+1}|}$$

- x^{k+1} estimación de x^* asociada a pesos actuales w^k:

$$x^{k+1} \in \underset{\substack{\Phi x = b \\ x \in \mathbb{R}^n}}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i|, \quad \forall k \geq 0$$

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Algoritmo RWℓ_1 de Candès, Wakin y Boyd [2008]

- Caso ideal $w_i = \frac{1}{|x_i^*|}$ sugiere actualizar pesos mediante:

$$w_i^{k+1} = \frac{1}{|x_i^{k+1}|}$$

- x^{k+1} estimación de x^* asociada a pesos actuales w^k:

$$x^{k+1} \in \underset{\substack{\Phi x = b \\ x \in \mathbb{R}^n}}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i|, \quad \forall k \geq 0$$

- En la práctica $w_i^{k+1} = \frac{1}{|x_i^{k+1}|+\epsilon}$, $\epsilon > 0$ (algoritmo CWB)

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Metodología RWℓ_1 propuesta

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

- Basada en relajación Lagrangeana
- Dada x^* solución esparsa, se define primal ideal:

$$\hat{x} \in \begin{array}{c} \text{argmin} \\ \Phi x = b \\ |x_i| \leq |x_i^*|,\ \forall\ i \end{array} \quad 0 \qquad (P)$$

- Como $\|\hat{x}\|_0 \leq \|x^*\|_0 \Rightarrow \hat{x}$ solución esparsa
- Problema convexo (fácil) pero ideal (requiere x^*)

Relajación restricciones ideales

- Relajando restricciones ideales $|x_i| \leq |x_i^*|$:

$$L(x,w) := 0 + \sum_{i=1}^{n} w_i \left(|x_i| - |x_i^*|\right) = \sum_{i=1}^{n} w_i |x_i| - \sum_{i=1}^{n} w_i |x_i^*|$$

- Función dual asociada:

$$d(w) := \min_{\substack{\Phi x = b \\ x \in \mathbb{R}^n}} L(w,x) =$$

$$= \left(\min_{\substack{\Phi x = b \\ x \in \mathbb{R}^n}} \sum_{i=1}^{n} w_i |x_i| \right) - \sum_{i=1}^{n} w_i |x_i^*|$$

- Problema "Weighted ℓ_1", w_i multiplicadores Lagrange

Problema
Minimización l_1
Problema Weighted ℓ_1
Algoritmo RWl_1 CWB
Metodología RWℓ_1 propuesta
Algoritmo RWℓ_1 subgradiente
Sparse Coding

Metodología propuesta para actualizar pesos

- Estimar w con algoritmos para estimar multiplicadores
- Alternativamente, estimar w como solución del Dual:

$$\begin{aligned} &\operatorname{argmax} \quad d(w) \qquad\qquad (D) \\ &w \geq 0 \\ &w \in \mathbb{R}^n \end{aligned}$$

- Maximización cóncava ($\Leftrightarrow$ minimización convexa)
- Función objetivo $d(w)$ no siempre diferenciable

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Algoritmo RWℓ_1 subgradiente

Solución del Dual con “Ascenso” por (sub)gradiente:

$$\begin{cases} w^{k+1} = w^k + \alpha_k g^k, \alpha_k > 0,\ g^k \in \partial d(w^k) \\ w^{k+1} = \max\{0, w^{k+1}\} \end{cases}$$

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Algoritmo RWℓ_1 subgradiente

Solución del Dual con “Ascenso” por (sub)gradiente:

$$\begin{cases} w^{k+1} = w^k + \alpha_k g^k, \alpha_k > 0,\ g^k \in \partial d(w^k) \\ w^{k+1} = \max\{0, w^{k+1}\} \end{cases}$$

Theorem (Subgradiente de función dual en w^k)

$$x^{k+1} \in \underset{\Phi x = b}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i| \Rightarrow$$

$$g(x^{k+1}) = |x^{k+1}| - |x^*| \in \partial d(w^k)$$

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Restricciones no Ideales

- Restricciones ideales son: $g_i(x) = |x_i| - |x_i^*|$
- Reemplazando x^* por estimación x^{k+1} ("amplificada"):

$$g_i^k(x) = |x_i| - (1 + \epsilon_k)\,|x_i^{k+1}|, \quad \epsilon_k > 0$$

Donde (como antes):

$$x^{k+1} \in \underset{\Phi x = b}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i|$$

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

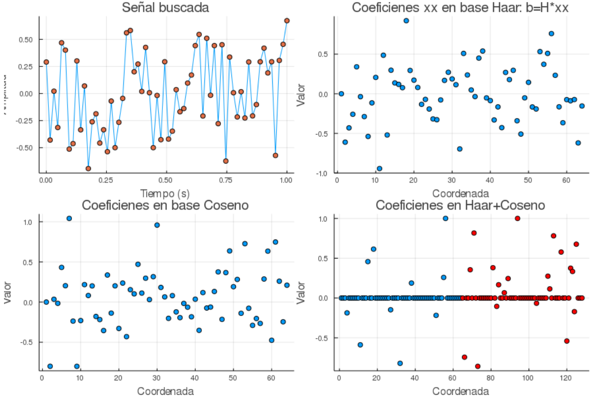
Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

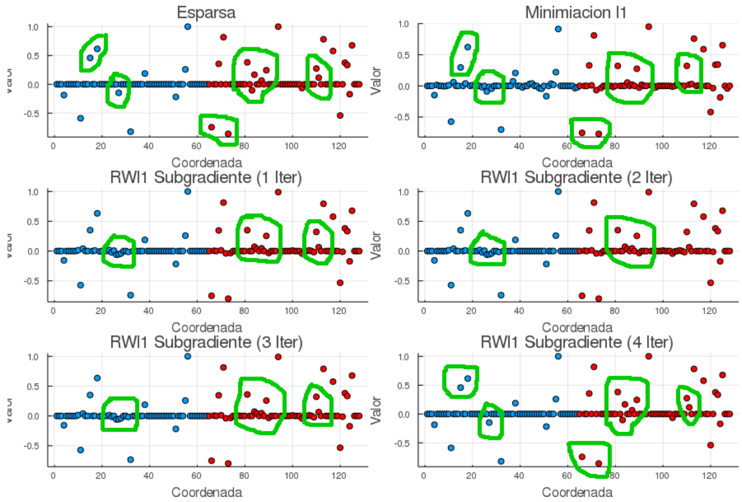
“Sparse Coding” (Representación Esparsa)

- Señal $b \in \mathbb{R}^n$: $b = \Phi x = \sum_{i=1}^{n} \Phi^i x_i$, para infinitos $x \in \mathbb{R}^n$
- $\Phi \in \mathbb{R}^{n \times 2n}$ Diccionario (unión de bases Haar y Coseno)
- Hallar representación más esparsa de b en Φ



Problema
Minimización l_1
Problema Weighted ℓ_1
Algoritmo RWl_1 CWB
Metodología RWℓ_1 propuesta
Algoritmo RWℓ_1 subgradiente
Sparse Coding

Sparse Coding: soluciones RWℓ_1 Subgradiente



Problema

Minimización l_1

Problema Weighted ℓ_1

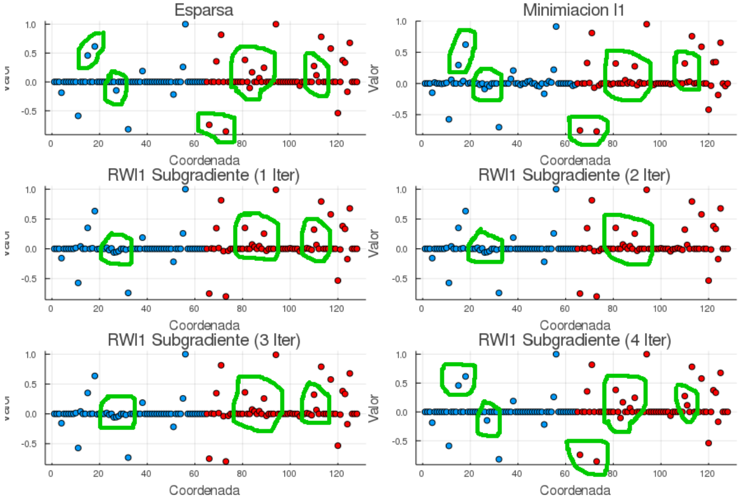
Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding

Sparse Coding: soluciones RWℓ_1 Subgradiente



Gitlab: `https://gitlab.fing.edu.uy/mvaldes/`

Problema

Minimización l_1

Problema Weighted ℓ_1

Algoritmo RWl_1 CWB

Metodología RWℓ_1 propuesta

Algoritmo RWℓ_1 subgradiente

Sparse Coding