L14: Orthogonal Matching Pursuit & Lasso and Compressed Sensing

Jeff M. Phillips

February 26, 2020

Mid Term

- No computers, calculators, phones
- 1 "cheat sheet" (both sides)
- Understand core definitions

3 Question (sub parts)

- Clustering
  - Assignment-based
  - HAC
- Similarity
- Distance
- Minhashing
- Jaccard
- k-grams
- Streaming
  MG

Ridge & Lasso Regression

Input $\tilde{X}, y$ $\quad X \in \mathbb{R}^{n \times d}$ $\quad y \in \mathbb{R}^n$ $\quad \tilde{X} = [1; X]$

Goal: RR: $\alpha_s^o = \underset{\alpha \in \mathbb{R}^{d+1}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \underbrace{\langle x_i, \alpha \rangle}_{M(x_i)})^2 + \underbrace{s\|\alpha\|_2^2}_{\text{regularize}}$

Lasso: $\alpha_s^\diamond = \underset{\alpha \in \mathbb{R}^{d+1}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \langle x_i, \alpha \rangle)^2 + s\|\alpha\|_1$

$$\alpha_t^o = \underset{\alpha \in \mathbb{R}^{d+1}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \langle x_i, \alpha \rangle)^2 \quad \text{s.t. } \|\alpha\|_2^2 \le t = \|\alpha_s^o\|_2^2$$

$$\alpha_t^\diamond = \underset{\alpha \in \mathbb{R}^{d+1}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \langle x_i, \alpha \rangle)^2 \quad \text{s.t. } \|\alpha\|_1 \le t$$

equivalent $\quad \forall$ choices $s \longrightarrow \exists t$

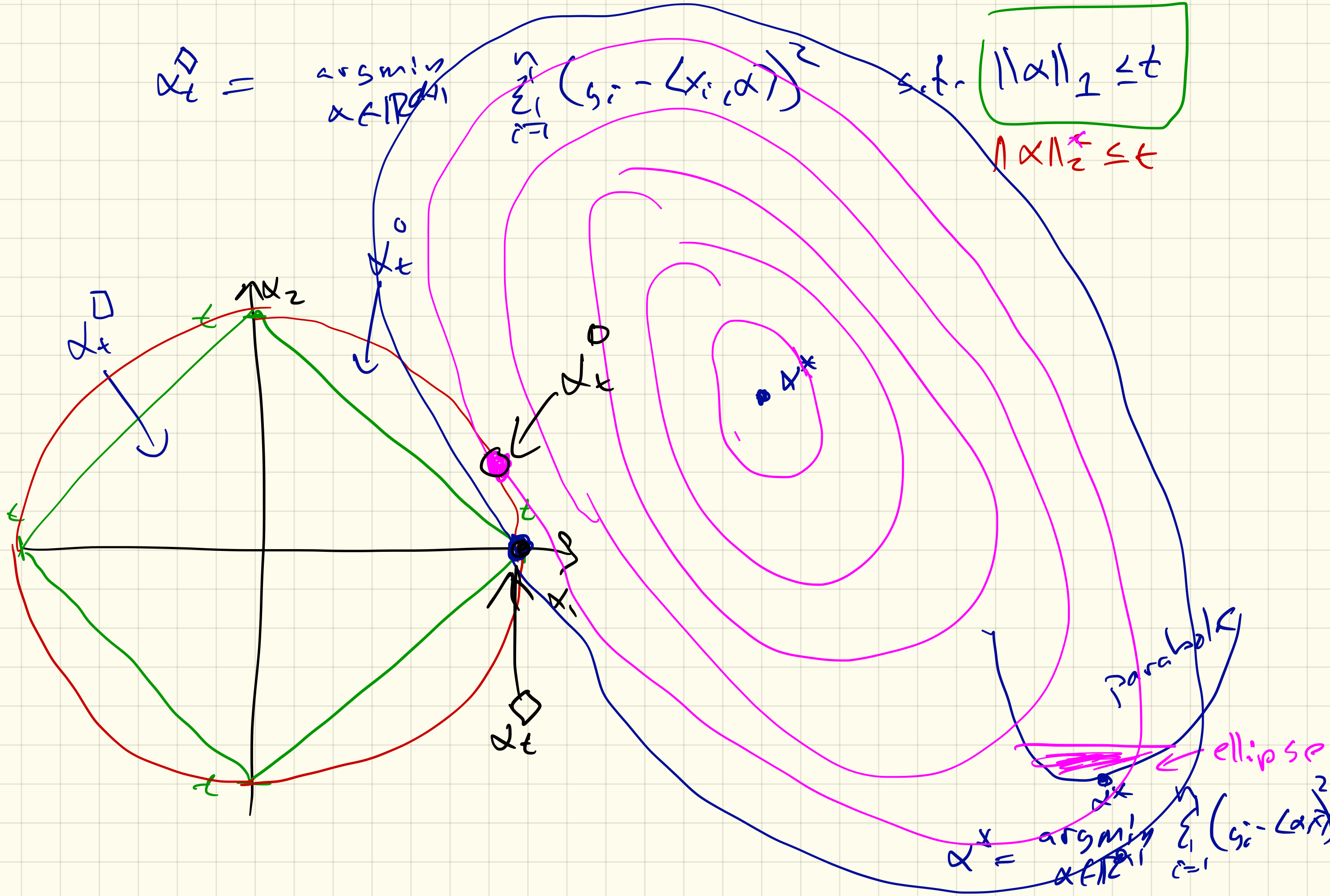
so $\alpha_s^o = \alpha_t^o$ (RR)

or $\alpha_s^\diamond = \alpha_t^\diamond$ (Lasso)

Let $t = \|\alpha_s^o\|_2^2$

$$\alpha_t^{\diamond} = \underset{\alpha \in \mathbb{R}^{d\times 1}}{\arg\min} \sum_{i=1}^{n} \left(y_i - \langle x_i, \alpha\rangle\right)^2 \quad \text{s.t.} \quad \boxed{\|\alpha\|_1 \le t}$$

$$\|\alpha\|_2^2 \le t$$

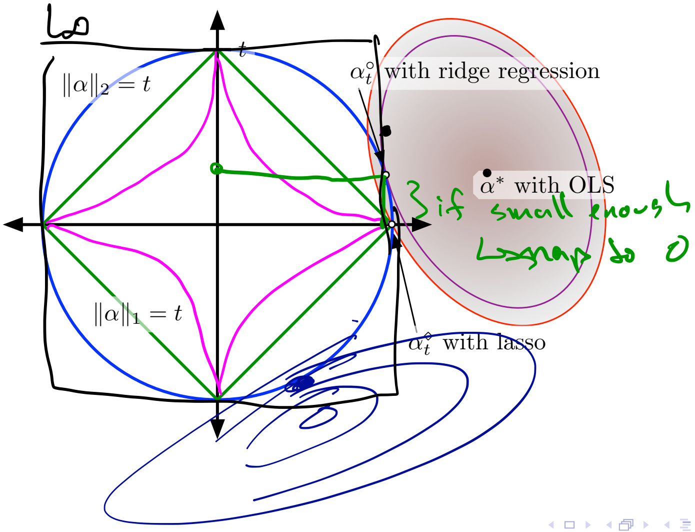


$$\alpha^* = \underset{\alpha \in \mathbb{R}^{d\times 1}}{\arg\min} \sum_{i=1}^{n} \left(y_i - \langle \alpha, x\rangle\right)^2$$

Lasso Illustration

$\|\alpha\|_{1/2}$

Find $\alpha^* = \arg\min_{\alpha \in \mathbb{R}^d} \|X\alpha - y\|_2 + s\|\alpha\|_1$



Matching Pursuit (MP)

Find $\alpha^* = \arg\min_{\alpha\in\mathbb{R}^d} \|X\alpha - y\|_2 + s\|\alpha\|_1$

Forward Subset Selection:

residual

<u>Matching Pursuit</u>

Set $r = y$, $\alpha = \mathbf{0}$.

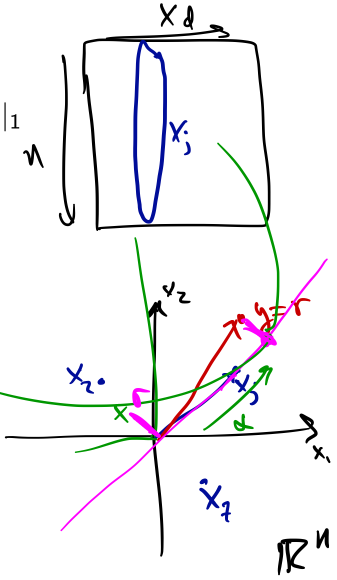
for $i = 1$ **to** t **do**

Set $X_j = \arg\max_{X_{j'} \in X} |\langle r, X_{j'} \rangle|$.

Set $\alpha_j = \arg\min_{\alpha} \|r - X_j\alpha\| + s|\alpha|$.

Set $r = r - X_j\alpha_j$.

Return α.



Orthogonal Matching Pursuit (OMP)

Find $\alpha^* = \arg\min_{\alpha \in \mathbb{R}^d} \|X\alpha - y\|_2 + s\|\alpha\|_2^2$
Forward Subset Selection:

Ridge

<u>Orthogonal Matching Pursuit</u>

Set $r = y$; $\alpha = \mathbf{0}$.
for $i = 1$ **to** t **do**
Set $X_j = \arg\max_{X_{j'} \in X} |\langle r, X_{j'} \rangle|$.
Set $\alpha = \arg\min_\alpha \|r - [X_1; X_2; \ldots; X_j]\alpha\| + s\|\alpha\|_2^2$.
Set $r = r - X_j\alpha_j$. (Update using other $\alpha_{j'}$ for $j' < j$)
Return α.

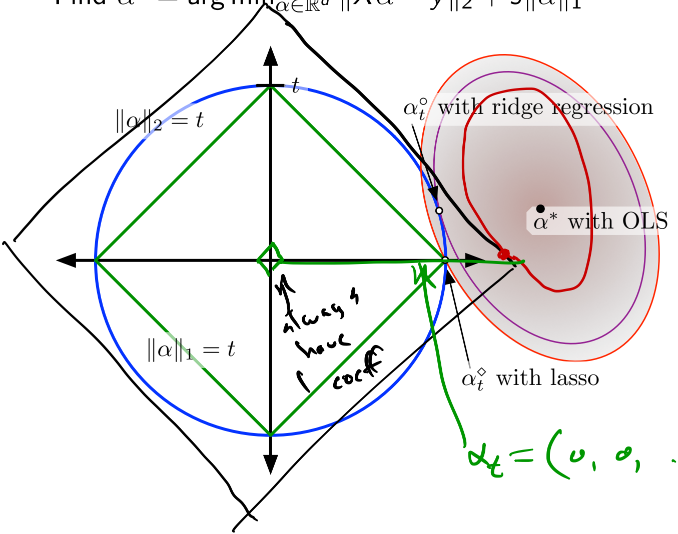
full soln on j coefficients

Lasso Illustration

Least Angle Regression

MP while increasing t

Find $\alpha^* = \arg\min_{\alpha \in \mathbb{R}^d} \|X\alpha - y\|_2 + s\|\alpha\|_1$



$\alpha_t = (0, 0, \ldots, \overset{j^*}{t}, 0, 0)$

Sparse Sensing

What if $d > n$

but $k > n$

$k \approx c\, n \log \frac{n}{k}$ (unknown)

$c \approx (2, 20)$

k sparse

signal d-dim bit vector

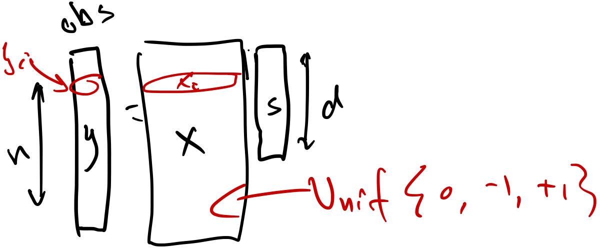
$a \approx$

$$S^T = [0\,1\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,1\,1\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0]$$

$$x_i^T = [\text{-}1\,0\,1\,0\,1\,1\,\text{-}1\,1\,0\,\text{-}1\,0\,0\,1\,\text{-}1\,\text{-}1\,1\,0\,1\,0\,1\,\text{-}1\,\text{-}1\,\text{-}1\,0\,1\,0\,0\,\text{-}1\,0\,1\,0\,0]$$

$$\begin{aligned} y_i &= \langle S, x_i \rangle \\ &= 0{+}0{+}0{+}0{+}1{+}0{+}0{+}0{+}0{+}0{+}0{+}0{+}0{+}0{+}0{+}1{+}0{+}0{+}0{+}1{-}1{+}0{-}1{+}0{+}0{+}0{+}0{+}0{+}0{+}1{+}0{+}0 \\ &= 2 \end{aligned}$$

observation i



Matching Pursuit (OMP)

Matching Pursuit

Set $r = y$.
for $i = 1$ **to** t **do**
Set $X_j = \arg\max_{X_{j'} \in X} |\langle r, X_{j'} \rangle|$.
Set $\gamma_j = \arg\min_\gamma \|r - X_j \gamma\|$.
Set $r = r - X_j \gamma_j$.
Return $\hat{S}$ where $\hat{s}_j = \gamma_j$ (or 0).

OMP Example

signal: $S = [0, 0, 1, 0, 0, 1, 0, 0, 1, 0]$

measurement: $X = \begin{bmatrix} 0 & 1 & 1 & -1 & -1 & 0 & -1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & 1 \\ 1 & -1 & 1 & -1 & 0 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 1 & -1 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & -1 & -1 & -1 & 0 & -1 & 1 & -1 & 0 \end{bmatrix}$

observation: $y = XS^T = [0, 0, 0, 1, 1, -2]^T$

$r_1 = [-1, 0, 0, 0, 0, -1]$