EE16A

Prof. Chunlei Liu

05/02/2019

To do:

Gram-Schmidt process

1. Motivation

2. steps

3. application O.M.P.

<u>Reading: application of "sparsity".</u>

Google scholar search:

Sparse MRI, M. Lustig et al. Magnetic Resonance in Medicine, 2007

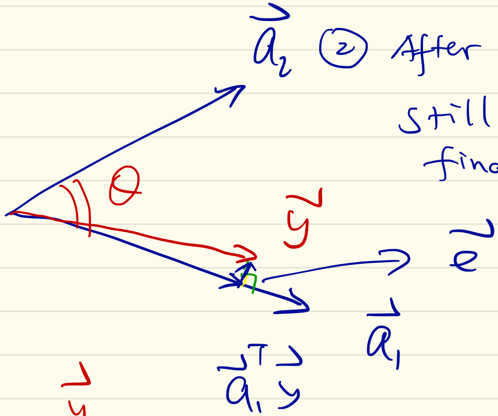
OMP Intuition

$$A\vec{x} = \vec{y}$$

$$\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$$

① If finding one code only, the closest one is $\vec{a}_1$

② After finding $\vec{a}_1$, there's still residual error, find next.



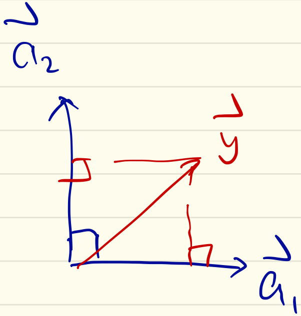
* generally $\vec{a}_1$, $\vec{a}_2$ not orthogonal.

$n = 2$

* if $\vec{a}_1 \perp \vec{a}_2$

much easier.

Just do projections.



Motivation

1. L.S. $\quad \hat{\vec{x}} = (A^T A)^{-1} A^T \vec{y}$

$$A^T A \in R^{n \times n} \sim O(n^3)$$

2. O.M.P. $\quad \hat{\vec{x}} = (A^T A)^{-1} A^T \vec{y}$

$$\vec{e} = \vec{y} - A\hat{\vec{x}} = \vec{y} - \underline{A(A^T A)^{-1} A^T \vec{y}}$$

In general, computation grows as # of iterations increases

* exception:

if columns of A are orthogonal, $A(A^T A)^{-1} A^T \vec{y} \rightarrow$ simple.

Let's see how ....

Given $A\vec{x} = \vec{y}$

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}, \qquad a_i^T \vec{a}_j = 0, \quad i \neq j$$

"orthogonal"

$$A^T A = \begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \\ & \vdots & \\ - & \vec{a}_n^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} I ? \end{bmatrix}, \quad \text{diagnolized?}$$

$$\checkmark \quad = \begin{bmatrix} \vec{a}_1^T \vec{a}_1 & & 0 \\ & \vec{a}_2^T \vec{a}_2 & \\ & & \ddots & \\ 0 & & & \vec{a}_n^T \vec{a}_n \end{bmatrix} ?$$

$$= \begin{bmatrix} \|\vec{a}_1\|_2^2 & & 0 \\ & \|\vec{a}_2\|_2^2 & \\ 0 & \ddots & \|\vec{a}_n\|_2^2 \end{bmatrix}$$

$$(A^TA)^{-1} = \begin{bmatrix} \frac{1}{\|\vec{a}_1\|_2^2} & & & 0 \\ & \frac{1}{\|\vec{a}_2\|_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\|\vec{a}_n\|_2^2} \end{bmatrix};$$

$$\underbrace{A\,(A^TA)^{-1}A^T\vec{y}}$$

$$A^T\vec{y} = \begin{bmatrix} \vec{a}_1^T\vec{y} \\ \vec{a}_2^T\vec{y} \\ \vdots \\ \vec{a}_n^T\vec{y} \end{bmatrix}$$

$$(A^TA)^{-1}A^T\vec{y} = \begin{bmatrix} \frac{\vec{a}_1^T\vec{y}}{\|\vec{a}_1\|_2^2} \\ \frac{\vec{a}_2^T\vec{y}}{\|\vec{a}_2\|_2^2} \\ \vdots \\ \frac{\vec{a}_n^T\vec{y}}{\|\vec{a}_n\|_2^2} \end{bmatrix}$$

recall $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$

$$\begin{bmatrix} \frac{\vec{a}_1^T\vec{y}}{\|a_1\|_2^2} & - & - & - \end{bmatrix}$$

next page

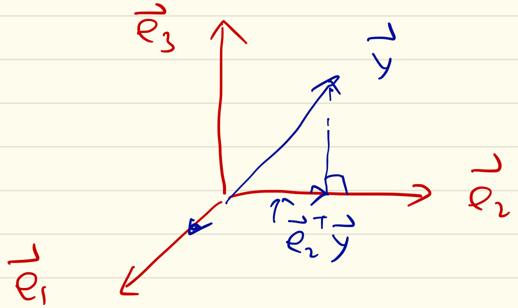
Linear combination of columns of A

$$A(A^T A)^{-1} A^T \vec{y}$$

$$= \frac{\vec{a}_1^T \vec{y}}{\|\vec{a}_1\|^2} \vec{a}_1 + \frac{\vec{a}_2^T \vec{y}}{\|\vec{a}_2\|^2} \vec{a}_2 + \cdots + \frac{\vec{a}_n^T \vec{y}}{\|\vec{a}_n\|^2} \vec{a}_n$$

define $\dfrac{\vec{a}_1}{\|\vec{a}_1\|} = \vec{e}_1$, $\|\vec{e}_1\| = 1$

$$= \underbrace{(\vec{e}_1^T \vec{y})}_{\triangle} \vec{e}_1 + (\vec{e}_2^T \vec{y}) \vec{e}_2 + \cdots + (\vec{e}_n^T \vec{y}) \vec{e}_n$$



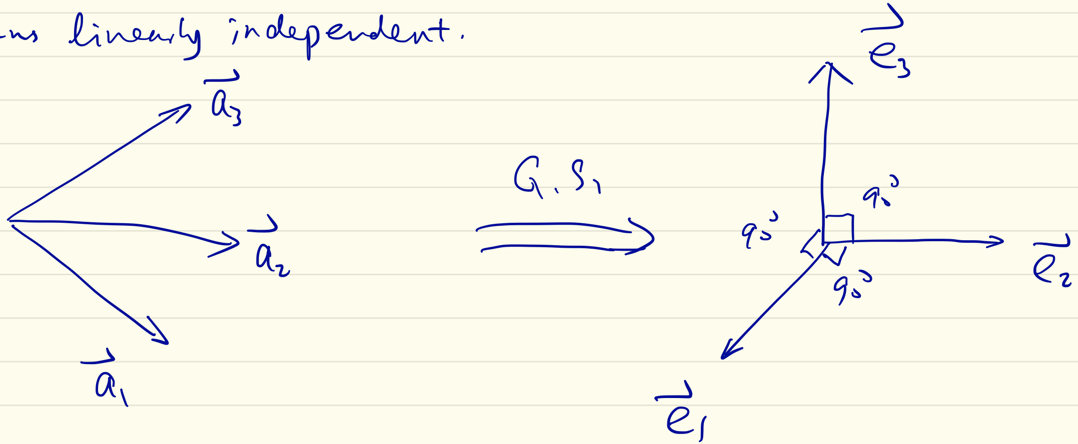
No matrix inversion needed !!

Simply projection.

* Gram Schmidt Process

$A \in R^{m \times n}$, columns are not orthogonal

columns linearly independent.



$$\text{span}(\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}) = \text{span}(\{\vec{e}_1, \vec{e}_2, \vec{e}_3\})$$

* property: $\vec{v}_1, \vec{v}_2$ are linearly independent

$$\operatorname{span}(\vec{v}_1, \vec{v}_2)$$

$$= \operatorname{span}(\vec{v}_1, \alpha \vec{v}_2) \quad , \quad \alpha \text{ is a "scalar"}$$

$$= \operatorname{span}(\vec{v}_1, \vec{v}_1 + \vec{v}_2)$$

$$= \operatorname{span}(\vec{v}_1, \vec{v}_1 - \vec{v}_2) = \operatorname{span}(\vec{v}_1, \vec{v}_2 - \vec{v}_1)$$

$$= \operatorname{span}(\vec{v}_1, \vec{v}_2 - \alpha \vec{v}_1)$$

Can I find an α, such that $\vec{v}_1 \perp (\vec{v}_2 - \alpha \vec{v}_1)$?

or $\underline{\vec{v}_1^{\,T} (\vec{v}_2 - \alpha \vec{v}_1) = 0}$

$$\Rightarrow \vec{v}_1^{\,T} \vec{v}_2 = \alpha \vec{v}_1^{\,T} \vec{v}_1 = \alpha \|\vec{v}_1\|_2^2$$

$$\Rightarrow \alpha = \frac{\vec{v}_1^{\,T} \vec{v}_2}{\|\vec{v}_1\|_2^2} \quad , \quad \vec{v}_2' = \vec{v}_2 - \alpha \vec{v}_1 = \vec{v}_2 - \frac{\vec{v}_1^{\,T} \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1$$

$$\frac{\vec{v}_1}{\|\vec{v}_1\|_2} = \vec{e}_1 \qquad = \vec{v}_2 - (\vec{e}_1^{\,T} \vec{v}_2)\, \vec{e}_1$$

$$\vec{e}_2 = \frac{\vec{v}_2'}{\|\vec{v}_2\|}$$

$$\text{span}(\vec{e}_1, \vec{e}_2) = \text{span}(\vec{v}_1, \vec{v}_2)$$

G.S.

step 1: Compute $\vec{e}_1 = \dfrac{\vec{a}_1}{\|\vec{a}_1\|}$, unit vector

step 2: for $i = 2, \cdots, n$

1) compute $\vec{q}_i = \vec{a}_i - \sum_{j=1}^{i-1} (\vec{e}_j^{\,T} \vec{a}_i)\, \vec{e}_j$

2) normalize $\vec{e}_i = \dfrac{\vec{q}_i}{\|\vec{q}_i\|}$

end.

$\Rightarrow \text{Span}(\{\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\}) = \text{Span}(\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\})$

OMP Summary

problem: $\vec{y} = x_1 \vec{S}_1^{(\tau_1)} + x_2 \vec{S}_2^{(\tau_2)} + \cdots + x_n \vec{S}_n^{(\tau_n)}$

$\vec{y} \in R^m$, $\vec{S}_i \in R^m$

also know: only k of $\{x_1, x_2, \cdots, x_n\}$ are non-zero

OMP procedures with Gram Schmidt

* set $\vec{e} = \vec{y}$, $j = 1$, $A = [\ ]$, $F = \{\phi\}$

$\hookrightarrow Q = [\ ]$

* while $[(j \le k)$ & $(\|\vec{e}\| \ge \varepsilon)]$

1. circorr$(\vec{e}, \vec{S}_i)$ for all $i \in \{1, 2, \cdots, n\}$,
   $\Longrightarrow$ Find which i^* such that circorr$(\vec{e}, \vec{S}_{i^*})$ has largest peak correlation.
2. Add i^* to $F = \{F, i^*\}$,
3. Add $\vec{S}_{i^*}^{(\tau_{i^*})}$ to A, $A = [A \mid \vec{S}_{i^*}^{(\tau_{i^*})}]$, as a new column vector

G.S. $\hookrightarrow \vec{e}_j$ $\qquad$ $Q = [Q \mid \vec{e}_j]$

4. Estimate a new sol. $\hat{\vec{x}} = \underbrace{(A^TA)^{-1}A^T}\vec{y}$, $[\underbrace{\vec{e}_1^T\vec{y}, \vec{e}_2^T\vec{y}, \cdots, \vec{e}_j^T\vec{y}}]^T$

$\hat{\vec{x}} \leftarrow [\hat{\vec{x}} \quad \vec{e}_j^T\vec{y}]^T$

5. Update residual $\vec{e} = \underline{\vec{y} - A\hat{\vec{x}}}$

$\vec{e}$ is not $\vec{e}_i'$

$\vec{y} - \left[\underbrace{(\vec{e}_1^T\vec{y})\vec{e}_1 + (\vec{e}_2^T\vec{y})\vec{e}_2 + \cdots} + (\vec{e}_j^T\vec{y})\vec{e}_j\right]$

$\vec{e} \leftarrow \vec{e} - (\vec{e}_j^T\vec{y})\vec{e}_j$

6. $j = j+1$,

* stop if $j > k$, or $\|\vec{e}\| < \varepsilon$

- With Gram Schmidt, the O.M.P. solves $A\vec{x} = \vec{y}$ in an orthonormal basis, denote as $\hat{\vec{x}}'$
- To compute $\hat{\vec{x}}$ for the original codes $\vec{S}_1^{(\tau_1)}, \vec{S}_2^{(\tau_2)}, \cdots, \vec{S}_n^{(\tau_n)}$, we need to perform a change of basis, using the relations btw $\hat{\vec{e}}_i$ & $\vec{S}_i^{(\tau_i)}$ defined by Gram-Schmidt.