Paul Hand

CAAM 654

Day 1

12 Jan 2015

- Linear Algebra
- Least Squares
- Intro to CS

Face recognition

$\{\phi_i, l_i\}_{i=1}^N$ — $\phi_i \in \mathbb{R}^m$ is image of person $l_i \in \{1, 2, \cdots, C\}$ under some illumination.

Given multiple images per person, and a new image $y \in \mathbb{R}^m$, determine which person y is.

Expect $y \approx \sum_i c_i \phi_i$ where c is nonzero only on images of single subject.

Let $\Phi = \begin{pmatrix} | & | & & | \\ \phi_1 & \phi_2 & \cdots & \phi_N \\ | & | & & | \end{pmatrix}$

Find c such that $y \approx \Phi c$ & c is mostly zero

c is sparse.

Sparse Image recovery

Consider a 100×100 pixel image of sky at night.
Most pixels black. Say, 10 pixels white.

How many measurements must be taken in order to recover image?

A measurement of $X \in \mathbb{R}^n$ is given by $a_i \cdot X = b_i$ for known a_i, b_i

Q: Is it $\approx 10{,}000$? or roughly 10? A: roughly 10

Potential: Acquire signal in compressed form, instead of measuring all pixels & throwing away data in compression.

Big Questions:

Setup: A signal $X \in \mathbb{R}^n$ has some sparsity structure.
Given m measurements of $X \in \mathbb{R}^n$, find X.

Qs: How many m are required?
How many m permit efficient algorithm for recovery?
What algorithm?
What types of measurements are ok?
What happens if there is noise?

Signal recovery without structure:

Let $x_0 \in \mathbb{R}^n$ be arbitrary. Let $a_i \cdot x_0 = b_i$ $\quad i = 1 \cdots m$

Given $\{a_i, b_i\}_{i=1}^m$ find x_0.

How many m are required?

$$m \geq n$$

Consider case $m = n$. How do you find x_0?

$$A = \begin{pmatrix} - & a_1 & - \\ - & a_2 & - \\ & \vdots & \\ - & a_n & - \end{pmatrix} \in \mathbb{R}^{n \times n}$$

Solve $Ax = b$ by LU and back substitution

When does this work?

If $\{a_i\}_{i=1}^n$ are linearly independent,

then x_0 is unique solution to $Ax = b$.

How robust to errors?

Suppose $\quad b = Ax_0 + e.$

Solve $\quad Ax = b.$

How big is error in x in terms of error in b?

$$A(x - x_0) = e$$

$$x - x_0 = A^{-1} e$$

spectral norm

$$\|x - x_0\|_2 \leq \|A^{-1}\| \, \|e\|_2$$

$$\|x - x_0\|_2 \leq \frac{1}{\sigma_{min}(A)} \|e\|_2$$

Relative error:

$$\cancel{\|x} \quad Ax_0 = b \quad \Rightarrow \quad \cancel{\|A} \quad \|b\| \le \|A\| \, \|x_0\|$$

So

$$\frac{\|x - x_0\|_2}{\|x_0\|_2} \le \|A\| \, \|A^{-1}\| \, \frac{\|\epsilon\|_2}{\|b\|_2}$$

$$\frac{\|x - x_0\|_2}{\|x_0\|_2} \le \underbrace{\frac{\sigma_{max}(A)}{\sigma_{min}(A)}}_{K} \, \frac{\|\epsilon\|_2}{\|b\|_2}$$

K - condition number

What if $m > n$?

Let $b_i = a_i \cdot x_0 \quad i = 1 \dots m \geq n, \quad a_i \in \mathbb{R}^n, \; x_0 \in \mathbb{R}^n$

Find x by solving least squares

$$\min_x \|Ax - b\|_2^2 \qquad \text{w/ } A = \begin{pmatrix} - a_1 - \\ - a_2 - \\ \vdots \\ - a_m - \end{pmatrix}$$

~~Home~~

How to solve:

~~Lagrange Multiplier~~

Take deriv & set to 0

$$\nabla \tfrac{1}{2} \|Ax - b\|_2^2 = \nabla \tfrac{1}{2} (Ax-b, Ax-b)$$

$$= A^t (Ax - b)$$

So $A^t A x = A^t b$

$$x = (A^t A)^{-1} A^t b$$

If $A = \hat{Q}\hat{R}$ is reduced QR

$$A = \hat{Q}\hat{R}$$

$$\underbrace{(\quad)}_{m \times n} = \underbrace{(\quad)}_{m \times n} \underbrace{(\quad)}_{n \times n}$$

$$x = (\hat{R}^t \hat{R})^{-1} \hat{R}^t \hat{Q}^t b$$

$$x = R^{-1} \hat{Q}^t b$$

Claim: If A is $m \times n$, $m > n$, full rank,
then $\arg\min_x \|Ax - b\|$ is x_0.

If $b = Ax + \epsilon$, then

$$\frac{\|x - x_0\|}{\|x_0\|} \leq K(\hat{R}) \frac{\|\epsilon\|_2}{\|b\|}$$ where $A = \hat{Q}\hat{R}$.

What measurement ~~matrices~~ vectors lead to better error bands?

If a_i are all orthogonal and normal, $\cancel{\sigma_{min}}\ K = 1$

If a_i are nearly parallel, $\cancel{\sigma_{min}}\ K$ ~~small~~ large

Compressed Sensing

~~Rk~~

Setup: $X_0 \in \mathbb{R}^n$ unknown signal
$\|x_0\|_0 = S \ll n$, so X_0 is sparse
measurement $a_i \cdot X_0 = b_i$ or $a_i \cdot x_0 = b_i + \epsilon_i$ noise
Given a_i & b_i, find X.

Want:

$$\min \|X\|_0 \text{ such that } Ax = b \quad \text{w/} \quad A = \begin{pmatrix} -a_1- \\ -a_2- \\ \vdots \\ -a_m- \end{pmatrix}$$

Can't minimize $\|x\|_0$ b/c NP-hard (combinatorially expensive)

Instead

$$\min \|X\|_1 \text{ such that } Ax = b$$

~~Thm:~~ If