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 $a_i \cdot x_0 = b_i$ $i = 1 \dots m$

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 $b = Ax_0 + e$.

Solve $Ax = b$.

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:

$$A x_0 = b \quad \Rightarrow \quad \|b\| \le \|A\| \, \|x_0\|$$

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

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

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, $\cancel{\sigma_{min}}\ K$ ~~≈ 0~~ large

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

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

~~How~~

How to solve:

~~Lagrange Multiplier~~

Take deriv & set to 0

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

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

$\langle A(x+\delta x) - b, A(x+\delta x) - b \rangle$

$= \langle Ax-b, Ax-b \rangle + \langle Ax-b, A\delta x \rangle$

$+ \langle A\delta x, Ax-b \rangle$

$+ o(\delta x^2)$

$= \langle Ax-b, Ax-b \rangle + \langle A^t(Ax-b), \delta x \rangle$

$+ \dots$

So

$A^t(Ax-b)$

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

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

$$\underset{m \times n}{\Big(\quad\Big)} = \underset{m \times n}{\Big(\quad\Big)} \underset{n \times n}{\Big(\quad\Big)}$$

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

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

Claim: If A is $m\times n$, $m>n$, full rank, $\underset{o}{b}=Ax_o$

then $\arg\min_x \|Ax-\underset{o}{b}\|$ is x_o.

If $b=Ax_o+\epsilon$, then

$$\frac{\|x-x_o\|_2}{\|x_o\|_2} \lesssim K(\hat{R}) \frac{\|\epsilon\|_2}{\|\underset{o}{b}\|_2}$$ where $A=\hat{Q}\hat{R}$.