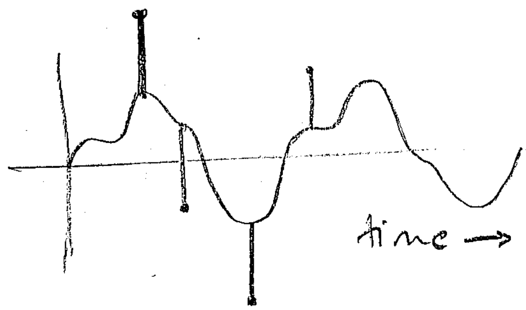
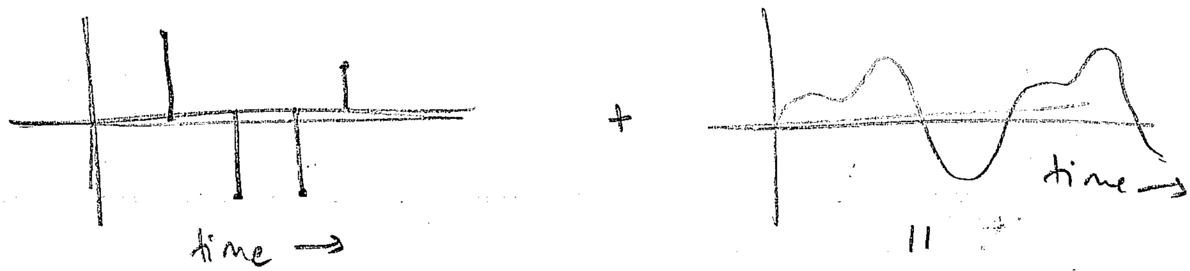
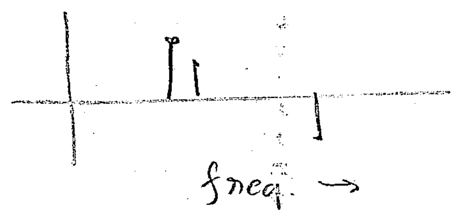
Sparse Representation in Pair of Bases

Paper by Elad & Bruckstein.

Supp. we are given a signal







ie addition of a sparse in time signal
w/ a " " freq. "

↳ to give one that is sparse in <u>neither</u>.

<u>Q</u>: Given a signal of this form, (when) can we ~~d.~~ find a sparse repn. in the "overcomplete" basis?

We will look at :- 2 bases Φ & Ψ.

<u>Basic issue</u>:- ill-posedness / non-uniqueness

↳ there are ∞ repn.s of any signal in a pair of bases.

looking for sparsity may help, but <u>not enough</u>.

eg:- if one basis is Φ, & Ψ is Φ w/ two vectors different so that both orthonormal

Sparse in Φ $\equiv$ sparse in Ψ as well.

hard for even a sparse x to have unique splitting.

☆ We need Φ & Ψ to be "different enough".

s.t. a vector cannot be sparse in both.

Basic Uncertainty Principle

Let Φ & Ψ be 2 orthonormal bases. For any vector $\underline{s} \neq 0$, let $\underline{s} = \Phi \underline{\alpha}$ & $\underline{s} = \Psi \underline{\beta}$. Let $M = \sup_{i,j} |\langle \phi_i, \psi_j \rangle|$

$\underline{\alpha}$: repr. in Φ $\quad$ $\underline{\beta}$: repr. in Ψ $\quad$ M: highest inner prod

Then $$\frac{\|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0}{2} \geq \sqrt{\|\underline{\alpha}\|_0 \|\underline{\beta}\|_0} \geq \frac{1}{M}$$ cannot be sparse in both.

Proof:- wlog $\underline{s}^T \underline{s} = 1$. Then

$$1 = \underline{\alpha}^T \Phi^T \Psi \underline{\beta} \leq \sum_i \sum_j |\alpha_i| |\langle \phi_i, \psi_j \rangle| |\beta_j| \leq M \left(\sum_i |\alpha_i|\right)\left(\sum_j |\beta_j|\right)$$

Let $A = \|\underline{\alpha}\|_0$ & $B = \|\underline{\beta}\|_0$

also $\sum_i |\alpha_i|^2 = 1$. $\qquad \sum_j |\beta_j|^2 = 1$

$\Rightarrow \sum_i |\alpha_i| \leq \sqrt{A} \qquad \sum_j |\beta_j| \leq \sqrt{B}$

(Begin w/ last
(First 2 pgs from 3/29 handwritten notes)

Corollary

Incoherence gives us uniqueness for the sparsest representation (ie any other repr. will have more non-zeros)

Thm If $\underline{s} = [\Phi\ \Psi]\, \underline{\gamma}_1$ & also $\underline{s} = [\Phi\ \Psi]\, \underline{\gamma}_2$ then

$$\|\underline{\gamma}_1\|_0 + \|\underline{\gamma}_2\|_0 \geq \frac{2}{M}.$$

Proof:- $\boxed{\text{HW!}}$

Uniqueness of sparse repn.

Thus :- $s = [\Phi, \Psi]\, \underline{\gamma}$ & $\|\underline{\gamma}\|_0 < \frac{1}{M}$, then there cannot be another $\tilde{\underline{\gamma}}$ s.t. $\underline{s} = [\Phi\ \Psi]\, \tilde{\underline{\gamma}}$ & $\|\tilde{\underline{\gamma}}\|_0 < \frac{1}{M}$.

Q: If $\underline{s}$ has unique sp. repn., how do we find it?

A:- $\min \|\underline{\gamma}\|_1$ s.t. $\underline{s} = [\Phi\ \Psi]\, \underline{\gamma}$

Now:- when does this work?

A: [Donoho & Huo] when $\|\underline{\gamma}\|_0 < \frac{1}{2}\left(1 + \frac{1}{M}\right)$

[Elad & Bruckstein] $\|\underline{\gamma}\|_0 < \frac{\sqrt{2} - 1/2}{M} \simeq \frac{0.9142}{M}.$

We will do [DH] result as proved in Sec. IV of [EB].

FOR

To show:- If $\|\underline{\gamma}\|_0 < F \Rightarrow \|\tilde{\underline{\gamma}}\|_1 \geq \|\gamma\|_1$ for any other $\tilde{\gamma}$ s.t. $\underline{s} = [\Phi\ \Psi]\tilde{\gamma}$.

now, $[\Phi\ \Psi](\tilde{\underline{\gamma}} - \underline{\gamma}) = 0$ Let $\underline{x} = \tilde{\underline{\gamma}} - \underline{\gamma}$. Thus need to show that

$$\sum_{k=1}^{2N} |\gamma_k + x_k| - \sum_{k=1}^{2N} |\gamma_k| \geq 0 \quad \forall\ x \text{ s.t. } [\Phi\ \Psi]\underline{x} = 0$$

$$\Leftrightarrow \underbrace{\sum_{k \in \mathrm{supp}(\underline{\gamma})}}_{\equiv \text{on}} \left(|\gamma_k + x_k| - |\gamma_k|\right) + \underbrace{\sum_{k \notin \mathrm{supp}(\gamma)}}_{\equiv \text{off}} |x_k| \geq 0.$$

$$\Leftarrow \sum_{\text{off}} |x_k| + \sum_{\text{on}} (-|x_k|) \geq 0$$ because $|a+b| - |a| \geq -|b|$.

$$\Leftrightarrow \sum_{\text{off}} |x_k| \geq \sum_{\text{on}} |x_k| \quad \Leftrightarrow \quad \frac{\sum_{\text{on}} |x_k|}{\sum_{\text{all}} |x_k|} \leq \frac{1}{2} \qquad \text{———} (*)$$

so suff. to prove this. now only a statement abt. support {ie no $\underline{\gamma}, \underline{s}$}

1. Let $\underline{x} = \begin{bmatrix} \underline{x}^{\Phi} \\ \underline{x}^{\Psi} \end{bmatrix}$ s.t. $\Phi \underline{x}^{\Phi} = -\Psi \underline{x}^{\Psi}$ $\Rightarrow x^{\Phi} = -\Phi^T \Psi x^{\Psi}$

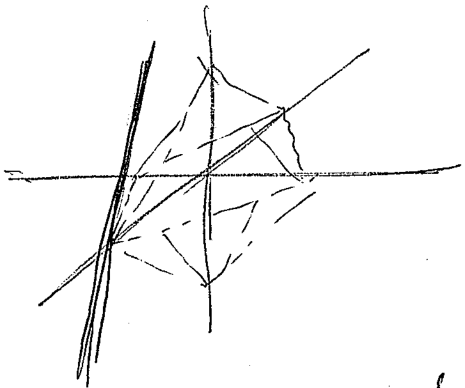
Suppose some co-ord. k has $x_k = V \neq 0$. Then
say in Φ

$$|V| = |x_k| = |\underbrace{[\Phi^T \Psi]_{k^{\text{th}}\text{ row}}}_{\text{every element of this row has abs. value} \leq M} \cdot x^{\Psi}| \leq M \|x^{\Psi}\|_1$$

also, $\|x^{\Phi}\|_1 \geq |x_k| = |V|$ so $\|x^{\Phi}\|_1 + \|x^{\Psi}\|_1 \geq |V|\left(1 + \frac{1}{M}\right)$

denominator of (*). if $x_k = V$

Pictures:- The set of γ s.t. $s = [\Phi\ \Psi]\gamma$ form a subspace of $\mathbb{R}^{2N}$. We want to find sparsest pnt in this subsp., (& cond's under which this is unique



3 dim's w/ 1-dim. subspace that has a unique 1-sparse pnt.

~~depending~~ for some of these subspaces, l_1-min. will find it & for others it won't.

subsp. $= \gamma_0 + a$ whe $[\Phi\ \Psi]a = 0$ ie a in null space.

Back to proof:

we want to see how large $\dfrac{\sum_{ON} |x_k|}{\sum_{ALL} |x_k|}$ can be (as a func. of $|ON|$)

for any x s.t. $[\Phi\ \Psi]x = 0$.

~~idea: let's min. denominator s.t. x is~~

Cons.

$\cancel{f_k(V) = \min \|x^\Phi\|_1 + \|x^\Psi\|_1 \text{ s.t. } \Phi x^\Phi = -\Psi x^\Psi \text{ \& } x_k}$

$\cancel{f_k(V) = \min \|x\|_1 \text{ s.t. } [\Phi\ \Psi]x = 0 \text{ \& } x_k = V}$

fix value to k^{th} coordinate

k in either "Φ part" or "Ψ part"

~~$f(V) = \min_k f_k(V)$~~

~~smallest denom. s.t. some co-ord = V.~~

~~s.t. $[\Phi\ \Psi]\underline{x} = 0$~~

Cons. any $\underline{x}$ & say that k is s.t. $|x_k| \geq |x_j|\ \forall j \neq k$.

Can rescale $\underline{x}$ s.t. $x_k = V$, w/o changing $\dfrac{\sum_{OW} |x_i|}{\sum_{ALL} |x_i|}$

now, for this (rescaled) x, $\sum_{OW} |x_i| \leq |V| \cdot |\mathrm{supp}(x)|$

& $\sum_{ALL} |x_i| \geq |V| \left(1 + \frac{1}{M}\right)$

$\therefore$ for any $\underline{x}$ st $[\Phi\ \Psi]\underline{x} = 0$, $\dfrac{\sum_{OW} |x_i|}{\sum_{ALL} |x_i|} \leq \dfrac{|\mathrm{supp.}(x)|}{1 + 1/M}$.

Thus, if $|\mathrm{supp}(x)| \equiv \|\underline{\gamma}\|_0 < \frac{1}{2}\left(1 + \frac{1}{M}\right)$

the unique sparse repr. of $\underline{s}$

Then we are done.

* This is an alternative proof technique (as opp. to what we saw for LASSO) that shows sparsity-inducing properties of l_1-min.

~~Compressed Sensing~~

~~Another~~

Comments:- • sp. repr. in pair of bases only works if they are <u>incoherent</u> $\{$ ie $M = \max_{i,j} \langle \phi_i, \psi_j \rangle$ is small$\}$.

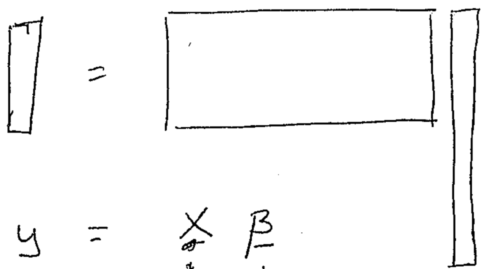
incoherence + sparsity = many things that seem "impossible" or "ill-defined" become "possible"

we saw:- • "unique" repr. in terms of dependent vectors

now:- "unique" solution to lin. eq.s when there are more vars than eq.s.

Compressed Sensing

Basic problem :-



in general:- an ∞ # of β's satisfy this.

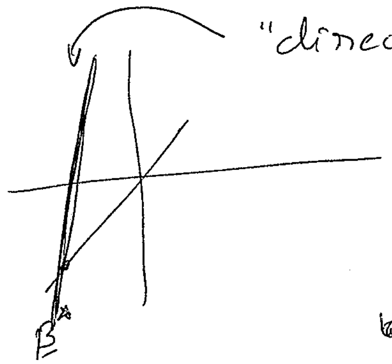
$y = X \beta$

y, X → known; β → to find.

<u>BUT</u> (under certain cond.s on X), there will be a unique <u>sparsest</u> soln.

IF your "truth" is this sparsest vector, some hope of finding it.

<u>picture</u> :- again similar



"direction" ie feasible subspace determined by X.

some subspaces "ball" &

(ie ℓ_1 will not find β^*).

we need to impose cond's on X s.t.

X is <u>always good for any β^* that is</u> <u>suff. sparse</u>.

<u>Background</u> :- CS vs classical compression

<u>CLASSICAL</u>

signal → Encoder → coefficients → Decoder → signal.

Encoder: complex. eg:- Fourier/wavelet transform.

Decoder: simple. eg:- linear combo. of known funcs.

<u>COMP. SENSE</u>

signal → Encoder → measurements → Decoder → signal.

Encoder: simple. eg:- linear maps of signal.

Decoder: complex. eg:- nonlinear optim.

↓ if decoder already knows rows of X.

Q: what cond's do we need on X s.t. sparsest vector is unique?

Supp. we have non-uniqueness for s-sparse vectors. ie,

A:- If we have 2 vectors $\underline{\beta}_1$ & $\underline{\beta}_2$ s.t. $|supp.(\underline{\beta}_1)| \leq s$

7

& $|supp(\underline{\beta}_2)| \leq s$ s.t. $X\underline{\beta}_1 = X\underline{\beta}_2$

$\Rightarrow X(\underbrace{\underline{\beta}_1 - \underline{\beta}_2}) = 0$

supp. of this is $\leq 2s$.

up to s

So:- for uniqueness up to s we need any 2s (or fewer) columns of X to be linearly independent.

Defn:- given some s, let δ_s = smallest number s.t.

$$(1-\delta_s)\|\underline{\beta}\|_2 \leq \|X\underline{\beta}\|_2 \leq (1+\delta_s)\|\underline{\beta}\|_2$$

for all $\underline{\beta}$ s.t. $|supp(\underline{\beta})| \leq s$

any s-sparse signal can be recovered uniquely (by any algorithm) $\iff \delta_{2s} < 1$.