E9 203 Compressed Sensing and
Sparse Signal Processing

26 Feb. 2021.

Last time:
- Course introduction
- Some key ideas

$\|x\|_0 = |\{j : x_j \neq 0\}|$

$A \in \mathbb{C}^{m\times N}$, $y \in \mathbb{C}^m$ $\quad \boxed{m, N, s}$

(P_0) $\boxed{\min_x \|x\|_0 \text{ s.t. } \boxed{Ax = y}}$
  - NP hard (combinatorial complexity)
  - Not robust to noise

$m \ll N$

s, k: sparsity of "true" x.

- Breakthroughs in CS:
  - Uniqueness: $\mathcal{N}(A)$ should not have sparse vecs. $\boxed{m \geq s+1}$.
  - Relaxation: ℓ_1 min. gives the same soln. as ℓ_0 min.
    - Measurement matrix is random
    - $\boxed{m = O\left(s \log\left(\frac{N}{s}\right)\right)}$ where s is the # nonzeros in the sparse x that resulted in y.

(P_1) $\boxed{\min \|x\|_1 \text{ s.t. } Ax = y}$ CONVEX.

  - Algorithms: MANY!
  - Guarantees.
- Key issues:
  - Computationally efficient algos
  - Robustness to noise
  - Stability if x is not exactly sparse.

Today: A closer look at underdetermined linear systems.

$$\boxed{Ax = y}, \quad A \in \mathbb{R}^{m\times N}, \quad m < N$$

No soln if $\text{rank}([y \,\vdots\, A]) > \text{rank}(A)$.

We will assume A is a full rank matrix.

Then $Ax = y$ has infinitely many solns.

$J(x)$: Undesirability of x as a soln. to $Ax = y$.

$$(P_J): \boxed{\min_{x\in\mathbb{R}^N} J(x) \quad \text{s.t. } Ax = y}$$

Consider $J(x) = \|x\|_2^2$

Lagrangian $L(x) = \|x\|_2^2 + \lambda^T(Ax - y)$

$$\frac{\partial L}{\partial x} = 2x + A^T\lambda = 0$$

$$\Rightarrow x_{opt} = -\frac{1}{2}A^T\lambda$$

$$Ax_{opt} = -\frac{1}{2}AA^T\lambda = y \Rightarrow \lambda = -2\underbrace{(AA^T)}_{m\times m}{}^{-1}y$$

$$\Rightarrow x_{opt} = A^T(AA^T)^{-1}y = A^\dagger y.$$

Closed-form and unique.

(P_J) convex if (a) $J(x)$ is a convex fn.
(b) Constraint set (feasible set) $[Ax = y]$ is a convex set.

Defn. (Convex Set): A set Ω is convex if $\forall x_1, x_2 \in \Omega$, $\forall t \in [0,1]$, convex comb. $x = tx_1 + (1-t)x_2 \in \Omega$.

Defn. (Convex fn.) A fn. $J(x): \Omega \to \mathbb{R}$ is convex if $\forall x_1, x_2 \in \Omega$ & $\forall t \in [0,1]$, $x = tx_1 + (1-t)x_2$ satisfies

$$J(tx_1 + (1-t)x_2) \leq t J(x_1) + (1-t) J(x_2).$$

If $J(\cdot)$ is twice continuously differentiable,

Defn. $J(x)$ is convex iff
(a) $\forall x_1, x_2 \in \Omega$, $J(x_2) \geq J(x_1) + (\nabla J(x_1))^T(x_2 - x_1)$
(b) $\forall x_1 \in \Omega$, $\nabla^2 J(x_1)$ (Hessian matrix) is positive semidefinite.

ℓ_2 norm square $\|x\|_2^2$ is convex $\because \nabla^2 \|x\|_2^2 = 2I > 0$.

Strictly convex $\Rightarrow$ min $J(x)$ for $J(x) = \|x\|_2^2$ has a unique soln.

$J(x) = \|Bx\|_2^2$, B nonsingular $N\times N$ matrix.

Special cases of interest for $J(x)$:

Defn. ℓ_p norm, $p \geq 1$: $\|x\|_p^p = \sum_{i=1}^N |x_i|^p$

ℓ_∞ norm: maximum magnitude entry of x

$\boxed{\ell_1 \text{ norm}}$: Sum of the absolute values of the entries in x.

ℓ_1-norm minimization:

$J(x) = \|x\|_1$ Convex fn, but not strictly convex

$t \in [0,1]$, $\|tx_1 + (1-t)x_2\| \leq t\|x_1\| + (1-t)\|x_2\|$.

Not strictly convex: x_1, x_2 same sign pattern

$$\|tx_1 + (1-t)x_2\|_1 = t\|x_1\|_1 + (1-t)\|x_2\|_1.$$

$$(P_1) \quad \min_{x\in\mathbb{R}^N} \|x\|_1 \quad \text{s.t. } \boxed{Ax = y}$$

This problem can have multiple solns. But:
1. The set of solns. form a bounded and convex set
2. $\exists$ at least 1 soln. with at most m nonzeros. $m \ll N$.

$A = [a_1\ a_2 \cdots a_N]$.

Take $A'_{m\times m} = [a_1\ a_2 \cdots a_m]$, set $x_{m+1} = \cdots = x_N = 0$.

$\underbrace{A'}_{m\times m} x' = \underbrace{y}_{m\times 1}$, where $x' = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$

if full rank, invertible and $x' = (A')^{-1}y$.

Boundedness: A & y are bounded, A full rank $\Rightarrow$
$V_{min} = \|x_{opt}\|_1 < \infty$. $\quad \|x^1_{opt}\|_1 = \|x^2_{opt}\|_1$

x^1_{opt}, x^2_{opt} are two optimal solns $\Rightarrow$

$$\|x^1_{opt} - x^2_{opt}\|_1 \leq \|x^1_{opt}\|_1 + \|x^2_{opt}\|_1 = 2V_{min}.$$

All solns are nearby $\Rightarrow$ bounded set.

Convexity: x^1_{opt}, x^2_{opt} solns to (P_1), say. $t \in [0,1]$

$$\|t x^1_{opt} + (1-t)x^2_{opt}\|_1 \leq t\underbrace{\|x^1_{opt}\|_1}_{V_{min}} + (1-t)\underbrace{\|x^2_{opt}\|_1}_{V_{min}} \leq V_{min}$$

If $\|\cdot\|_1 < V_{min}$, $\exists$ a better soln. $\because$

$$A(\underbrace{t x^1_{opt} + (1-t)x^2_{opt}}_{\text{feasible pt.}}) = tAx^1_{opt} + (1-t)Ax^2_{opt} = ty + (1-t)y = y$$

$\to$ contradiction $\Rightarrow \|t x^1_{opt} + (1-t)x^2_{opt}\|_1 = V_{min}$.

$\Rightarrow$ Set of solns. form a convex set.

$\exists$ a soln. with at most m nonzeros:

Say x_{opt} is a soln. to (P_1) & has $k > m$ nonzeros.

$\exists\ 0 \neq h \in \mathbb{R}^N$ s.t. $\text{supp}(h) \subseteq \text{supp}(x_{opt})$ and $Ah = 0$

[For any $x \in \mathbb{R}^N$, $\text{supp}(x) \triangleq \{j : x_j \neq 0\}$.]

Consider $x = x_{opt} + \varepsilon h \Rightarrow Ax = Ax_{opt} + \varepsilon Ah = y$

Let $|\varepsilon| \leq \min_i \frac{|x_{opt,i}|}{|h_i|}$, $i \in \text{supp}(h)$

$\text{sign}(x) = \text{sign}(x_{opt})$ same sign pattern.

Since x_{opt} is an optimal soln,

$$\|x_{opt} + \varepsilon h\|_1 \geq \|x_{opt}\|_1 \quad — (*)$$

Since ε can be +ve or −ve, (*) can hold for all ε only if $\|x_{opt} + \varepsilon h\|_1 = \|x_{opt}\|_1 \Rightarrow$ $\boxed{x \text{ is also an optimal soln.}}$

Alternatively, $h^T \text{sign}(x_{opt}) = 0$.

Choose $\varepsilon = -\frac{x_{opt,j}}{h_j}$ where $j = \arg\min_{i\in\text{supp}(h)} \frac{|x_{opt,i}|}{|h_i|}$

In the resulting $x = x_{opt} + \varepsilon h$ has $x_j = 0$

and $\|x\|_1 = \|x_{opt}\|_1$.

Thus we have a new soln. x which solves (P_1) and has at most $(k-1)$ nonzero entries.

Can repeat until $k = m$.

$\Rightarrow (P_1)$ has an optimal soln. w/ at most m nonzeros.

ℓ_1 minimization and linear programming

(P_1) $\min_{x \in \mathbb{R}^N} \|x\|_1$ s.t. $Ax = y$

Let $x = \underline{u} - \underline{v}$, u = all +ve entries in x, 0 elsewhere

$v = -$(all −ve entries in x, 0 ")

$$\|x\|_1 = \|u\|_1 + \|v\|_1 = \mathbb{1}_N^T u + \mathbb{1}_N^T v = \mathbb{1}^T (u+v)$$
$$= \mathbb{1}_{2N}^T z, \quad \text{where } z = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^{2N}$$

Then, $Ax = A(u-v) = [A \ \ -A] \begin{bmatrix} u \\ v \end{bmatrix}$

$= [A \ -A] z = y$: constraint.

(LP) : $\boxed{\min_{z \in \mathbb{R}^{2N}} \mathbb{1}^T z \text{ s.t. } [A \ -A] z = y, \ z \geq 0.}$

LINEAR PROGRAM

Suppose k^{th} entry in both u & v obtained by solving the LP are nonzero. $u_k > v_k > 0$

Replace the k^{th} entries by $u_k' = u_k - v_k$ and $v_k' = 0$.

Satisfies $y = [A \ \ -A] z$, $z \geq 0$

$\|z\|_1$ is smaller the prev. soln! Reduced by $u_k - v_k$ which contradicts the optimality of the prev. soln.

$\Rightarrow$ the supports of u & v obtained by solving the LP do not overlap i.e., $(P_1) \equiv (LP)$.

$$u = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 0 \end{bmatrix}$$

$\text{Supp}(u) = \{1, 3\}$, $\text{supp}(v) = \{4\}$.

$[N] = \{\underline{1, 2, \ldots, N}\}$

$\text{Supp}(x) \subseteq [N]$ for any $x \in \mathbb{R}^N$