· l_1 norm heuristics for cardinality problems
(number of non-zeros)

Cardinality problems arise often, but are hard to solve exactly
a simple heuristic, that relies on l_1-norm, seems to work well
(sparse design, LASSO, SVM, total variation reconstruction in signal processing, compressed sensing)

· Cardinality

the cardinality of $x \in \mathbb{R}^n$, denoted $\mathrm{card}(x)$, is the number of nonzero components in x
card( ) is separable   for scalar x, $\mathrm{card}(x) = \begin{cases} 0 & x=0 \\ 1 & x \neq 0 \end{cases}$
card( ) is quasiconcave on $\mathbb{R}^n_+$ (but not on $\mathbb{R}^n$)
$\mathrm{card}(x+y) \geq \min\{\mathrm{card}(x), \mathrm{card}(y)\}$

· General convex-cardinality problem
a convex-cardinality problem is one that would be convex,
except for appearance of card( ) in objective and constraints

$$\begin{aligned} \min.\ & \mathrm{card}(x) \\ \text{s.t.}\ & x \in C \end{aligned} \qquad\qquad \begin{aligned} \min.\ & f(x) \\ \text{s.t.}\ & x \in C \quad \mathrm{card}(x) \leq k \end{aligned}$$

· Boolean LP as convex-cardinality problem

$$\begin{aligned} \min.\ & c^T x \\ \text{s.t.}\ & Ax \leq b \quad x_i \in \{0,1\} \end{aligned}$$

$$\Updownarrow$$

$$\begin{aligned} \min.\ & c^T x \\ \text{s.t.}\ & Ax \leq b \quad \mathrm{card}(x) + \mathrm{card}(1-x) \leq n \end{aligned}$$

$\mathrm{card}(x) + \mathrm{card}(1-x) \leq n \iff x_i \in \{0,1\}$

Sparse design: $\min.\ \mathrm{card}(x)$ s.t. $x \in C$

Sparse modeling / regressor selection: $\min.\ \|Ax-b\|_2$ s.t. $\mathrm{card}(x) \leq k$
select k features

· Sparse signal reconstruction

estimate signal x, given

noisy measurement $y=Ax+v$ $v\sim\mathcal{N}(0,\sigma^2 I)$, A is known, v unknown

prior information $\mathrm{card}(x)\le k$

maximum likelihood estimation

$$\begin{aligned} \min.\ & \|Ax-y\|_2 \\ \text{s.t.}\ & \mathrm{card}(x)\le k \end{aligned}$$

· Estimation with outliers

noisy measurment $y_i = a_i^T x + v_i + w_i$ $v_i \sim \mathcal{N}(0,\sigma^2)$

only assumption on w is sparsity $\mathrm{card}(w)\le k$ nonzero w_i yields outlier

maximum liklihood estimation

$$\begin{aligned} \min.\ & \sum_{i\notin B} (y_i - a_i^T x)^2 \\ \text{s.t.}\ & |B|\le k \end{aligned} \qquad B=\{i \mid w_i\neq 0\}$$

$$\Downarrow$$

$$\begin{aligned} \min.\ & \|y-Ax-w\|_2^2 \\ \text{s.t.}\ & \mathrm{card}(w)\le k \end{aligned}$$

· Minimum number of violations

$$\begin{aligned} \min.\ & \mathrm{card}(t) \\ \text{s.t.}\ & f_i(x)\le t_i \\ & x\in C \quad t_i\ge 0 \end{aligned}$$

· Linear classifier with fewest errors

given data $(x_1,y_1)\ \cdots\ (x_m,y_m)\in\mathbb{R}^n\times\{1,-1\}$

seek linear (affine) classifier $y\approx \mathrm{sign}(w^T x+v)$

classifier error : $y_i(w^T x_i + v)$

SVM !!!

$$\begin{aligned} \min.\ & \mathrm{card}(t) \\ \text{s.t.}\ & y_i(w^T x_i+v)+t_i\ge 1 \end{aligned}$$

we use homogeneity in w, v

· Smallest set of mutually infeasible inequalities

given a set of mutually infeasible convex inequalities

$f_1(x) \le 0 \quad \cdots \quad f_m(x) \le 0$

find smallest (cardinality) subset of these that is infeasible

Strong statement: inequality #14 alone is infeasible

less strong statement: #14, #22 and #37 are mutually infeasible

weakest statement: all inequalities are mutually infeasible,

if I remove anyone of them, the others becomes feasible

which one (ones) mess you up

certificate of infeasibility: $g(\lambda) = \inf_x \left(\sum_{i=1}^{m} \lambda_i f_i(x) \right) \ge 1 \qquad \lambda \succeq 0$

find smallest cardinality infeasible subset

$\min. \ \mathrm{card}(\lambda)$

$\text{s.t.} \ g(\lambda) \ge 1 \quad \lambda \succeq 0$

if $\lambda_i = 0$

you didn't use λ_i in constructing the feasibility

· Piecewise constant fitting

fit corrupted signal x_{cor} by a piecewise constant signal $\bar{x}$ with k or fewer jumps

$D = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & & \ddots & \\ & & & 1 & -1 \end{bmatrix}$ is the difference matrix

$\mathrm{card}(\underbrace{D\bar{x}}_{\text{difference}}) \le k$

convex cardinality problem:

$\min. \ \|x_{cor} - \bar{x}\|_2$

$\text{s.t.} \ \mathrm{card}(D\bar{x}) \le k$

· Piecewise linear fitting

$\min. \ \|\bar{x} - x_{cor}\|_2$

$\text{s.t.} \ \mathrm{card}(\underbrace{\nabla \bar{x}}_{\text{curvature}}) \le k$

$\nabla = \begin{bmatrix} -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & & \ddots & & \\ & & & -1 & 2 & -1 \end{bmatrix}$

· l_1-norm heuristic

replace $\mathrm{card}(z)$ with $\gamma \|z\|_1$, or add regularization term $\gamma\|z\|_1$ to objective

more sophisticated version:

$\sum_i w_i |z_i| \qquad \sum_i w_i (z_i)_+ + \sum_i v_i (z_i)_- \qquad (w_i, v_i > 0)$

- eg minimum cardinality problem

$$\begin{array}{ll} \min. & \mathbf{card}(x) \\ \text{s.t.} & x \in C \end{array} \quad \rightarrow \quad \begin{array}{ll} \min. & |x| \\ \text{s.t.} & x \in C \end{array}$$

eg. cardinality constrained problem

$$\begin{array}{ll} \min. & f(x) \\ \text{s.t.} & x \in C \quad \mathbf{card}(x) \le k \end{array}$$

$$\begin{array}{ll} \min. & f(x) \\ \text{s.t.} & x \in C \quad \|x\|_1 \le \beta \end{array} \qquad \begin{array}{ll} \min. & f(x) + \gamma\|x\|_1 \\ \text{s.t.} & x \in C \end{array} \quad \Bigg\} \text{ adjust } \beta, \gamma \text{ so that } \mathbf{card}(x) \le k$$

- Polishing

use a heuristic to find x with required sparsity

fix the sparsity pattern of x

resolve the (convex) problem with this sparsity pattern

eg. regressor selection

$$\begin{array}{ll} \min. & \|Ax-b\|_2 \\ \text{s.t.} & \mathbf{card}(x) \le k \end{array} \quad \rightarrow \quad \min. \ \|Ax-b\|_2 + \gamma\|x\|_1$$