400101204 کیمیا یاسری

HW3 - Compressed sensing

تمرین سری 3 مسائل فشرده

مسئله 1:

(I) $\sqrt{k} \cdot \max_{i \in T} \{ |\Phi_i \cdot y| \} \geq \| \Phi_T^T y \|_2 = \| \Phi_T^T \Phi_T x + \Phi_T^T v \|_2$

$|T| = k$, $y = \Phi_T x + v$

Φ_T is the columns in accordance with the x

using triangle identity.

$\Rightarrow \sqrt{k} \cdot \max_{i \in T} \{ |\Phi_i \cdot y| \} \geq \| \Phi_T^T \Phi_T x \|_2 - \| \Phi_T^T v \|_2$

$\geq \lambda_{min}(\Phi_T^T \Phi_T) \cdot \|x\|_2 - \lambda_{max}(\Phi_T^T \Phi_T) \cdot \|v\|_2$

$\geq (1-\delta_{k+1}) \cdot \|x\|_2 - \sqrt{1+\delta_{k+1}} \cdot \|v\|_2$ Q.E.D.

(II) since supp x & supp j don't have intersection, then

$\langle \Phi_{t'}, \Phi_T x \rangle = \langle \Phi_{t'}, \Phi_T x \rangle \leq \delta_{k+1} \|\delta\| \cdot \|x\| = \delta_{k+1} \cdot \|x\|_2^2$

$\Rightarrow |\langle \Phi_{t'}, y \rangle| = |\langle \Phi_{t'}, \Phi x \rangle + \langle \Phi_{t'}, v \rangle| \leq |\langle \Phi_{t'}, \Phi x \rangle| + |\langle \Phi_{t'}, v \rangle|$

$\leq \delta_{k+1} \cdot \|x\|_2 + \sqrt{1+\delta_1} \cdot \|v\|_2 < \delta_{k+1} \cdot \|x\|_2 + \sqrt{1+\delta_k} \cdot \|v\|_2$

$\sqrt{1+\delta_1} < \sqrt{1+\delta_k}$ Q.E.D.

(III) the first step of OMP. let $t' = \arg\max_{i \in \Omega} |\langle \Phi_i, y \rangle|$ and according to (I) we get:

$\max_{i \in \Omega} |\langle \Phi_i, y \rangle| > \max_{i \in T} |\langle \Phi_i, y \rangle| > \frac{1}{\sqrt{k}} \left[(1-\delta) \|x\|_2 + \sqrt{1+\delta} \|v\|_2 \right]$

if $t' \notin T$, then according to (II) we'd obtain:

$\delta_{k+1} \cdot \|x\|_2 + \sqrt{1+\delta_1} \cdot \|v\|_2 \geq |\langle \Phi_{t'}, y \rangle| > \frac{1}{\sqrt{k}} (1-\delta_{k+1}) \|x\|_2 - \frac{\sqrt{1+\delta_{k+1}}}{\sqrt{k}} \cdot \|v\|_2$

thus: $\left(\frac{1-\delta}{\sqrt{k}} - \delta \right) \cdot \|x\|_2 < \left(1 + \frac{1}{\sqrt{k}}\right) \sqrt{1+\delta} \cdot \|v\|_2 = \left(1 + \frac{1}{\sqrt{k}}\right) \sqrt{1+\delta} \, \frac{\|\Phi x\|_2}{\sqrt{SNR}}$

$\Rightarrow \left[\frac{1-\delta}{\sqrt{k}} - \delta \right] \|x\|_2 < \left(1 + \frac{1}{\sqrt{k}}\right) \sqrt{1+\delta} \, \frac{\|\Phi x\|_2}{\sqrt{SNR}} < \left(\frac{1-\delta}{\sqrt{k}} - \delta \right) \frac{\|\Phi x\|_2}{\sqrt{1+\delta}}$

which leads to (since $\|\Phi x\|_2 < \lambda_{max} \cdot \|x\|_2 = \sqrt{1+\delta} \cdot \|x\|_2$

$\rightarrow t'$ is surely in T ($t' \in T$), $\|x\| \leq \frac{\|\Phi x\|_2}{\sqrt{1+\delta}}$ Q.E.D.

$$z_{opt} = \underset{z}{\text{argmin}} \{ \overbrace{\|y - Az\|_2^2 + \lambda \|z\|_1}^{J(z)} \} \; 2\,O' \, 2$$

(I) we know that at the optimal point $z_{opt} \to \vec{0} \in \partial J(z_{opt})$

$$\frac{\partial J(z)}{\partial z_j} = -2(y - Az) \cdot A_j + \lambda . \text{sgn}(z) \xrightarrow{z=0} 2y . A_j + \gamma = 0$$

which suggests $\gamma = \partial J(\vec{z}) = [-\lambda, \lambda]$ is the subgradient.

thus λ must be such that $\max_j \{ |\langle y, A_j \rangle| \} = 0$

therefore, $\lambda > \max_j \{ 2A_j^T y \}$. $\square$ Q.E.D ■

(II) Assume that z_1, z_2 are two solutions. consider $\Delta = z_1 + \varepsilon (z_1 - z_2)$.

we take ε so small that the sign of the z_1 and Δ are the same.

thus: $\|\Delta\|_1 = \|z_1\|_1 + \varepsilon \underbrace{\sum_i (z_1 - z_2)_i}_{S}$. we take the sign of ε such that $\varepsilon . S < 0$.

thus: $\|y - A\Delta\|_2^2 = \|y - Az_1\|_2^2 + \varepsilon^2 \|A(z_1 - z_2)\|_2^2 + 2\varepsilon \langle y - Az_1, A(z_1 - z_2) \rangle$

(*) (1) $\leq \|y - Az_1\|_2^2 + \varepsilon^2 \|A(z_1 - z_2)\|_2^2 - \varepsilon \lambda \sum_i (z_1 - z_2)_i$

since $\nabla J(z_{opt}) = 0$, thus the vector $2(y - Az_1)^T A$ has elements in $[-\lambda, \lambda]$ and at most $\to \lambda . \text{sgn}(\varepsilon (z_1 - z_2))$

(a) if (*) is not true, then we get

$$\|y - A\Delta\|_2^2 < \|y - Az_1\|_2^2 + \lambda \|z_1\|_1 + \overbrace{\varepsilon^2 + \varepsilon \underbrace{(\text{---})}_{S_0}}^{T(\varepsilon)}$$

thus $\exists \varepsilon$ which is small and also $T(\varepsilon) < 0$ ■ Q.E.D

thus we got Δ which is less than z_1 and this is a contradiction $\times$.

(b) if the equality holds $\to 2(y - Az_1)^T A = \lambda \, \text{sgn}(z_1 - z_2)$. Likewise for $\vec{z}_2$ we get $2(y - Az_2)^T A = \lambda \, \text{sgn}(z_2 - z_1)$

$$\Rightarrow [A(z_1 - z_2)]^T A = 0 \Rightarrow A^T A z_1 = A^T A z_2 \Rightarrow A z_1 = A z_2$$

if $\lambda = 0$ still $\exists \varepsilon$: $|\varepsilon| \ll 1$ and $T(\varepsilon) < 0$, still leads to contradiction $\times$.

$Az_1 = Az_2$

(III) $\left.\begin{array}{l}\|y - Az_1\| + \lambda\|z_1\|_1 \\ \|y - Az_2\| + \lambda\|z_2\|_1\end{array}\right\} = \to \|z_1\|_1 = \|z_2\|_1$

(I) $1 \le n \le s$:3 [illegible]

lemma: if $\max_{j\in S} |\langle A_j, r\rangle| > \max_{\ell\in S^c} |\langle A_\ell, r\rangle| \quad \forall r \in \{Az \neq 0 : \operatorname{sup}(z) = s\}$ holds, then OMP at each step will return the indices inside S.

Proof: at the first step, since $\vec{y} = A\vec{x}$, $\operatorname{argmax}_j |\langle A_j, y\rangle|$ will return $j \in S$, thus $S' \subseteq S$. Now using induction we'll prove that theorem. Assume that $S^i \subseteq S$, and $r^i = y - Ax^i$: $\forall\, 1 \le i \le n$

since $x^n = A_n^+ y \to Ar^n = A_{S^n} r^n = 0$. so at the $n+1$-th step: $\operatorname{argmax}_j |\langle A_j, r^n\rangle|$ will return indices inside S. thus $S^{n+1} \subseteq S$.. & after s-steps it will return S. ✓ Q.E.D.

so it suffices to prove $\max_{j\in S} |\langle A_j, r\rangle| > \max_{\ell} |\langle A_\ell, r\rangle|$. let's write $\vec{r}$ as $\vec{r} = \sum_{i\in S} r_i A_i$. Therefore: $\left(\text{let } |r_k| = \max_{i\in S} |r_i|\right)$

$\ell \in S^c \to |\langle r, A_\ell\rangle| = \left|\sum_{i\in S} r_i (A_i^T A_\ell)\right| \le \mu_1(s) \cdot |r_k|$

$j \in S \to |\langle r, A_j\rangle| = \left|\sum_{i\in S} r_i (A_j^T A_i)\right| \ge |r_j| \langle A_j, A_j\rangle - \sum_{\substack{i \notin S \\ i \neq j}} |r_i| \langle a_i, a_k\rangle$

$\ge |r_k| - |r_k| \mu_1(s-1)$

Thus, if $\langle \vec{r}, \vec{A_k}\rangle > \langle \vec{r}, \vec{A_\ell}\rangle \Rightarrow \mu_1(s)|r_k| \ge |r_k|(1 - \mu_1(s-1))$

$\Rightarrow \boxed{\mu_1(s) + \mu_1(s-1) < 1}$ Q.E.D ■

(II) in the last HW(2) we proved that necessary & sufficient condition for that BP problem has a unique solution is that $\forall v \in N(\Phi) - \{\vec{0}\} \to \|v\|_{S,1} < \|v\|_{\bar{S},1}$. therefore, it only sufficies to prove that $\|v\|_{S,1} < \|v\|_{\bar{S},1}$

ادامه سوال 3.

let $v \in N(A) \setminus \{\vec{0}\} \Rightarrow \sum_i A_i \vec{v}_i = 0$

$\forall i \in S: \quad v_i = v_i \langle A_i, A_i \rangle - \sum_i A_i v_i = -\sum_{l \in \bar{S}} v_l \langle a_l, a_i \rangle - \sum_{\substack{j \neq i \\ j \in S}} v_j \langle a_j, a_i \rangle$

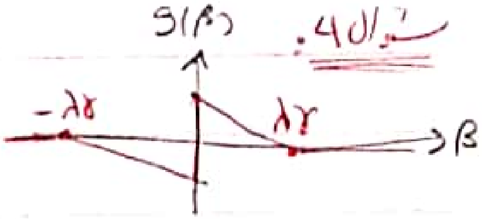
$$\Rightarrow \|\vec{v}\|_{S,1} = \sum_{i \in S} |v_i| < \sum_{i \in S} \Big[\sum_{l \in \bar{S}} |v_l| \, |\langle a_l, a_i \rangle| + \sum_{\substack{j \neq i \\ j \in S}} |v_j| \, |\langle a_j, a_i \rangle|$$

$$\Rightarrow \|\vec{v}\|_{S,1} < \|v_{\bar{S}}\|_1 \, \mu_1(S) + \|v_S\|_1 \, \mu_1(S-1)$$

$$\Rightarrow \|v_S\|_1 < \frac{\mu_1(S)}{1-\mu_1(S-1)} \|v_{\bar{S}}\|_1 < \|v_{\bar{S}}\|_1$$

thus, the BP(Basis Pursuit) algorithm will reconstruct the signal exactly.

سوال 4.



$$\min_{\beta} \left\{ \frac{1}{2}(\beta - \hat{\beta})^2 + \lambda \int^{|\beta|} \left(1 - \frac{x}{\lambda\gamma}\right)_+ dx \right\}$$

$\Rightarrow \vec{0} \in \frac{\partial f}{\partial \beta} = (\beta - \hat{\beta}) + g(\beta)$

the subgradient at $\beta = 0$ is $[-\lambda, \lambda]$. so we've three cases:

(i) $\hat{\beta} > \lambda\gamma \Rightarrow \frac{\partial f}{\partial \beta} = \beta - \hat{\beta} + g(\beta) = 0 \longrightarrow g(\hat{\beta}) = 0 \longrightarrow \beta = \hat{\beta}$

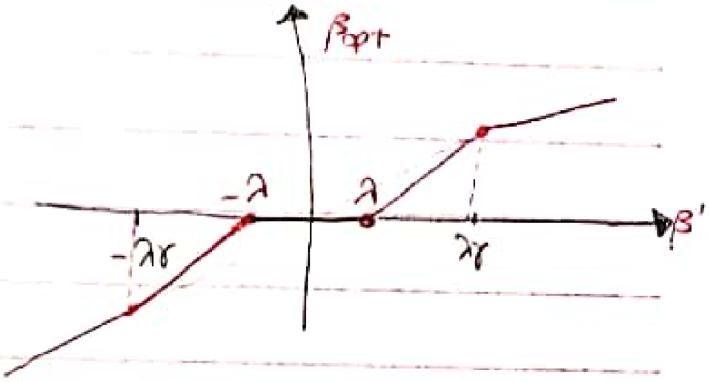
(ii) $\lambda < \hat{\beta} < \lambda\gamma$. Assume that solution lies in $[0, \lambda\gamma] \Rightarrow$

$\frac{\partial f}{\partial \beta} = \beta - \hat{\beta} + \lambda - \frac{\beta}{\gamma} = 0 \longrightarrow 0 < \beta \doteq \frac{\hat{\beta} - \lambda}{1 - 1/\gamma} < \lambda\gamma$. Thus the optimal solution lies in $[0, \lambda\gamma]$ which is correct ✓.

(iii) $0 < \hat{\beta} < \lambda$: in this case $\beta - \hat{\beta}$ will touch the vertical vertex within the interval $[-\lambda, \lambda] \Rightarrow \vec{0} \in \frac{\partial f}{\partial \beta}\Big|_{\beta=0} = [\underbrace{-\lambda - \hat{\beta}}_{<0}, \underbrace{\lambda - \hat{\beta}}_{>0}]$

السؤال 4 :

(iv) at the case $\hat{\beta} < 0$, since $\frac{\partial f}{\partial \beta}$ is an odd function wrt β, the same thing as in (i), (ii) & (iii) will happen as well.



Thus:
$$\beta_{opt} = \begin{cases} \hat{\beta} & : \ |\hat{\beta}| > \lambda\gamma \\ 0 & : \ |\hat{\beta}| < \lambda \\ \dfrac{\hat{\beta} - \lambda}{1 - \frac{1}{\gamma}} & : \ \lambda < \hat{\beta} < \lambda\gamma \\ \dfrac{\hat{\beta} + \lambda}{1 - \frac{1}{\gamma}} & \ -\lambda\gamma < \hat{\beta} < -\lambda \end{cases}$$

سؤال 5 :

In order to have full-reconstruction for k-sparse signals we must have $RIP(2k, \simeq \frac{1}{2})$, we know that $1 - \delta_{2k} = \min\{\lambda_{min}(\Phi_{2k}^T \Phi_{2k})\}$ if the vectors are normals according to Gershgorin Theorem, $\lambda_{min}(\Phi_{2k}^T \Phi_{2k})$ is at least $1 - \sum_{j \neq i} \langle a_i, a_j \rangle$. $\sum_{j \neq i} \langle a_i, a_j \rangle$ is at most $\mu_1 + \mu_2 + \dots + \mu_{2k}$. Since if $\mu_1 + \dots + \mu_{2k} < \frac{1}{2}$: the Reconstruction will be exact (ادامه)

The largest k such that $\frac{1}{2} > \sum_{i=1}^{2k} \mu_i$, so k-sparse vectors will have full reconstructions. $\frac{1}{2} > \mu_1 + \dots + \mu_{2k}$

on the other hand, according to Problem 2, if $\mu(s) + \mu(s-1) < 1$, then we have full reconstruction with BP & OMP algorithm.

we have $\begin{cases} \mu(s) < \mu_1 + \dots + \mu_s \\ \mu(s-1) < \mu_1 + \dots + \mu_{s-1} \end{cases}$. if $2\mu_1 + \dots + 2\mu_{k-1} + \mu_k < 1$, we've full reconstruction.

ادامه $\Rightarrow$ the largest k such that: $\frac{1}{2} > \mu_1 + \dots + \mu_{k-1} + \frac{\mu_k}{2}$, the k-sparse vectors will have full reconstructions which is more strong than other.

Q.E.D ∎

lemma:

:6 Ql3

i) if $a, b \in \Sigma_k$ and $supp(a) \cap supp(b) = \phi$, then let $\vec{\alpha} = \frac{\vec{a}}{\|a\|_2}, \vec{\beta} = \frac{\vec{b}}{\|b\|_2}$

$\Rightarrow (1-\delta_{2k}) \|\alpha \pm \beta\|_2^2 \le \|\Phi(\alpha \pm \beta)\|_2^2 \le (1+\delta_{2k}) \|\alpha \pm \beta\|_2^2$

$\Rightarrow -4\delta_{2k} \le \underbrace{\|\Phi(\alpha+\beta)\|_2^2 - \|\Phi(\alpha-\beta)\|_2^2}_{<\Phi\alpha, \Phi\beta>} \le 4\delta_{2k}$

which suggests $|<\Phi\vec{\alpha}, \Phi\vec{\beta}>| = \frac{|<\Phi\vec{a}, \Phi\vec{b}>|}{\|a\|_2 \cdot \|b\|_1} \le \delta_{2k}$

$\Rightarrow \boxed{|<\Phi\vec{a}, \Phi\vec{b}>| \le \delta_{2k} \cdot \|a\|_2 \cdot \|b\|_2}$ Q.E.D. ■

ii) $(1-\delta_{2k}) \|h_\Delta\|_2^2 \le \|\Phi h_\Delta\|_2^2 = <\Phi h_\Delta, \Phi h_\Delta> = <\Phi h_\Delta, \Phi h - \sum_{i\ge 2} \Phi h_{\Delta_i}>$

$= <\Phi \vec{h_\Delta}, \Phi \vec{h}> - \sum_{i \ge 2} <\Phi h_{\Delta_0}, \Phi h_{\Delta_i}> - \sum_{i\ge 2} <\Phi h_{\Delta_1}, \Phi h_{\Delta_i}>$

$\le <\Phi h_\Delta, \Phi h> - \sum_{i \ge 2} \delta_{2k} \left[\|h_{\Delta_0}\|_2 \cdot \|h_{\Delta_i}\|_2 + \|h_{\Delta_1}\|_2 \cdot \|h_{\Delta_i}\|_2 \right]$

$= <\Phi h_\Delta, \Phi h> + \delta_{2k} \left(\|h_{\Delta_0}\|_2 + \|h_{\Delta_1}\|_2 \right) \left(\sum_{i \ge 2} \|h_{\Delta_i}\|_2 \right)$

Now since $\|h_{\Delta_0}\|_2 + \|h_{\Delta_1}\|_2 \le \sqrt{2} \cdot \|h_\Delta\|_2$

and also $\frac{\|h_{\Delta_{i-1}}\|_2}{\sqrt{k}} \ge \|h_{\Delta_i}\|_1$ and the lemma above

will lead to:

$\|\Phi h_\Delta\|_2^2 \le <\Phi h_\Delta, \Phi h> + \sqrt{2}\, \delta_{2k} \cdot \|h_\Delta\|_2 \cdot \frac{\|h_{\Delta_0^c}\|_1}{\sqrt{k}}$

$\hookrightarrow \boxed{\|h_\Delta\|_2 \le \frac{\sqrt{2}\,\delta_{2k}}{1-\delta_{2k}} \cdot \frac{\|h_{\Delta_0^c}\|_1}{\sqrt{k}} + \frac{1}{1-\delta_{2k}} <\Phi h_\Delta, \Phi h>}$

Q.E.D. ■

(I) applying maximum likelihood:

$$\arg\max_x P(y|x) = \arg\max_x \left\{ (2\pi\sigma)^{-\frac{m}{2}} \cdot e^{-\frac{\|y-Ax\|_2^2}{2\sigma^2}} \right\}$$

take log and neglect constants wrt x

$$= \arg\min_x \left\{ \|y-Ax\|_2^2 \right\}$$

Q.E.D ■

(II) $$\arg\max_x P(y|x) = \arg\max_x \left\{ (2a)^{-m} e^{-\frac{\|y-Ax\|_1}{a}} \right\}$$

take log

$$= \arg\min_x \left\{ \|y-Ax\|_1 \right\}$$

Q.E.D. ■

(III) $$\arg\max_x p(y|x) = \arg\max_x \left\{ (2a)^{-m} \prod_i \mathbb{1}\left[|(y-Ax)_i| < a \right] \right\}$$

all must equal to 1

$\Rightarrow$ which means: $\forall i: |(y-Ax)_i| < a \rightarrow \boxed{\|y-Ax\|_\infty \le a}$

Q.E.D. ■

(IV) $$\arg\max_x P(y|x) = \arg\max_{a,x} \left\{ (2a)^{-m} \prod_i \mathbb{1}\left(|y-Ax|_i < a \right) \right\}$$

$$= \arg\min_{a,x} \; a \quad \text{s.t. } \|y-Ax\|_\infty \le a$$

Q.E.D. ■

$$= \arg\min_x \|y-Ax\|_\infty$$