CS 754 HW 2

Group: Piyush Bharambe (18D070019)
Vinit Awale (18D070067)

Q1)

Sol^n: (1) Justify how $\delta_{2S} = 1$ could imply that 2S columns of ϕ may be linearly dependent.

$\rightarrow$

Let $h = x - x'$ where x and x' are two distinct s-sparse vectors. Then h is $2s$-sparse

Using RIP,

$$(1 - \delta_{2s})\|h\|_2^2 \le \|\phi h\|_2^2 \le (1+\delta_{2s})\|h\|_2^2$$

Since $\delta_{2s} = 1$,

$$0 \le \|\phi h\|_2^2 \le 2\|h\|_2^2$$

$\Rightarrow \|\phi(x - x')\|_2^2 \geq 0$

Thus $\|\phi h\|_2^2$ may be 0 for some $h = x - x'$ (where $x \neq x'$). In that case, we have $\phi h = 0$. Since h is $2s$-sparse, there are $2s$ columns of ϕ for which $\sum_{i=1}^{2s} \phi^{(i)} h_i = 0$ ($\phi^{(i)}$ is i^{th} column & h_i is non-zero element of h)

Hence, for $\delta_{2s} = 1$, $2s$ columns of ϕ may be linearly dependent.

(2) Justify inequalities

$$\|\phi(x^* - x)\|_{\ell_2} \leq \|\phi x^* - y\|_{\ell_2} + \|y - \phi x\|_{\ell_2} \leq 2\epsilon$$

→ Consider,

$$\|\phi(x^* - x)\|_{\ell_2} = \|(\phi x^* - y) + (y - \phi x)\|_{\ell_2}$$

By triangle inequality,

$$\|(\phi x^* - y) + (y - \phi x)\|_{\ell_2} \leq \|\phi x^* - y\|_{\ell_2} + \|y - \phi x\|_{\ell_2} \quad \text{— (1)}$$

Also x & x^* satisfy the constraint of the optimization problem → $\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1$ s.t. $\|y - \phi\tilde{x}\|_2 \leq \epsilon$

Using this, we have,

$$\|\phi x^* - y\|_2 \leq \epsilon \text{ and } \|y - \phi x\|_2 \leq \epsilon$$

$$\therefore \|\phi x^* - y\|_2 + \|y - \phi x\|_2 \leq 2\epsilon \quad \text{— (2)}$$

Using (1) & (2), we conclude,

$$\|\phi(x^* - x)\|_{\ell_2} \leq \|\phi x^* - y\|_{\ell_2} + \|y - \phi x\|_{\ell_2} \leq 2\epsilon$$

(3) Justify inequality: $\|h_{T_j}\|_{\ell_2} \leq s^{1/2} \|h_{T_j}\|_{\ell_\infty} \leq s^{-1/2} \|h_{T_{j-1}}\|_{\ell_1}$

→ For each ~~h~~ j, h_{T_j} is s-sparse.

By definition, ℓ_∞ norm of vector is the maximum value element of that vector.

Now, $\|h_{T_j}\|_{\ell_2} = \sqrt{\sum_{i \in T_j} h_i^2} \leq \sqrt{s \times (\max_{i \in T_j}(h_i))^2}$

$$\therefore \|h_{T_j}\|_{\ell_2} \leq \sqrt{s}\, \|h_{T_j}\|_{\infty} \quad \text{— (3)}$$

By construction of vectors $\{h_{T_j}\}$, the largest element of h_{T_j} is smaller than the smallest element of $h_{T_{j-1}}$. So, the largest element of h_{T_j} is smaller than every element of $h_{T_{j-1}}$.

$$\therefore \|h_{T_j}\|_{\ell_\infty} \leq h_p \qquad p \in T_{j-1}$$

Summing for all such ~~elemes~~ s non zero elements.

$$s\|h_{T_j}\|_{\ell_\infty} \leq \sum_{p\in T_{j-1}} h_p \leq \sum_{p \in T_{j-1}} |h_p|$$

$$\therefore s\|h_{T_j}\|_{\ell_\infty} \leq \|h_{T_{j-1}}\|_{\ell_1}$$

$$\therefore s^{1/2}\|h_{T_j}\|_{\ell_\infty} \leq s^{-1/2}\|h_{T_{j-1}}\|_{\ell_1} \quad \text{---} \quad (4)$$

From ~~(1) & (2)~~, (3) & (4),

$\|h_{T_j}\|_{\ell_2} \leq s^{1/2}\|h_{T_j}\|_{\ell_\infty} \leq s^{-1/2}\|h_{T_{j-1}}\|_{\ell_1}$ is proved.

(4) Justify inequality:

$$\sum_{j\geq 2}\|h_{T_j}\|_{\ell_2} \leq s^{-1/2}(\|h_{T_1}\|_{\ell_1} + \|h_{T_2}\|_{\ell_1} + \dots) \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

(4) Justify inequality.

$$\sum_{j\geq 2} \|h_{T_j}\|_{\ell_2} \leq s^{-1/2}(\|h_{T_1}\|_{\ell_1} + \|h_{T_2}\|_{\ell_1} + \dots) \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

→ From previous part, we have,

$$\|h_{T_j}\|_{\ell_2} \leq s^{-1/2}\|h_{T_{j-1}}\|_{\ell_1}$$

Summing for all $j \geq 2$, we get,

$$\sum_{j\geq 2} \|h_{T_j}\|_{\ell_2} \leq s^{-1/2}(\|h_{T_1}\|_{\ell_1} + \|h_{T_2}\|_{\ell_1} + \dots\dots) \quad (5)$$

By construction of $\{h_{T_j}\}$, we have

$$\|h_{T_1}\|_{\ell_1} + \|h_{T_2}\|_{\ell_1} + \dots\dots = \|h_{T_0^c}\|_{\ell_1} \quad (6)$$

From (5), (6), we conclude

$$\sum_{j\geq 2} \|h_{T_j}\|_{\ell_2} \leq s^{-1/2}(\|h_{T_1}\|_{\ell_1} + \|h_{T_2}\|_{\ell_1} + \dots\dots) \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

(5) Justify:

$$\|h_{(T_0\cup T_1)^c}\|_{\ell_2} = \Big\|\sum_{j\geq 2} h_{T_j}\Big\|_{\ell_2} \leq \sum_{j\geq 2}\|h_{T_j}\|_{\ell_2} \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

→ By construction we have,

$$h_{(T_0\cup T_1)^c} = \sum_{j\geq 2} h_{T_j}$$

Thus, $\|h_{(T_0\cup T_1)^c}\|_{\ell_2} = \Big\|\sum_{j\geq 2} h_{T_j}\Big\|_{\ell_2}$

Using triangle inequality, $\Big\|\sum_{j\geq 2} h_{T_j}\Big\|_{\ell_2} \leq \sum_{j\geq 2}\|h_{T_j}\|_{\ell_2}$

From previous part, we have $\sum_{j\geq 2}\|h_{T_j}\|_{\ell_2} \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$

Combining above inequalities, we get

$$\|h_{(T_0\cup T_1)^c}\|_{\ell_2} = \Big\|\sum_{j\geq 2} h_{T_j}\Big\|_{\ell_2} \leq \sum_{j\geq 2}\|h_{T_j}\|_{\ell_2} \leq s^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

(6) Justify: $\sum_{i\in T_0}|x_i+h_i| + \sum_{i\in T_0^c}|x_i+h_i| \geq (\|x_{T_0}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} - \|x_{T_0^c}\|_{\ell_1})$

→ By the reverse triangle inequality, we have,

$$|x_i+h_i| \geq ||x_i|-|h_i|| \geq |x_i|-|h_i|$$

Summing over $i \in T_0$,

$$\sum_{i\in T_0}|x_i+h_i| \geq \sum_{i\in T_0}|x_i| - \sum_{i\in T_0}|h_i|$$

$$\therefore \sum_{i\in T_0}|x_i+h_i| \geq \|x_{T_0}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} \quad \text{— (7)}$$

Doing the same for $i \in T_0^c$, we get,

$$\sum_{i\in T_0^c}|x_i+h_i| \geq \|h_{T_0^c}\|_{\ell_1} - \|x_{T_0^c}\|_{\ell_1} \quad \text{— (8)}$$

Adding (8) & (7), we get required inequality

(7) Justify: $\|h_{T_0^c}\|_{\ell_1} \leq \|h_{T_0}\|_{\ell_1} + 2\|x_{T_0^c}\|_{\ell_1}$

→ From previous part, we have,

$$\|x\|_{\ell_1} \geq \|x_{T_0}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} - \|x_{T_0^c}\|_{\ell_1} \quad \text{— (9)}$$

By definition, $\|x_{T_0^c}\|_{\ell_1} = \|x - x_S\|_{\ell_1}$

& $x_S = x_{T_0}$.

Using reverse triangle inequality,

$$\|x_{T_0^c}\|_{\ell_1} = \|x-x_S\|_{\ell_1} \geq \|x\|_{\ell_1} - \|x_S\|_{\ell_1} = \|x\|_1 - \|x_{T_0}\|_{\ell_1} \quad \text{— (10)}$$

From (9) & (10), we get,

$$2\|x_{T_0^c}\|_{\ell_1} + \|h_{T_0}\|_{\ell_1} \geq \|h_{T_0^c}\|_{\ell_1}$$

↳ hence proved.

(8) Justify: $\|h_{(T_0 \cup T_1)^c}\|_{\ell_2} \le \|h_{T_0}\|_{\ell_2} + 2e_0$; $e_0 = \frac{\|x - x_s\|_{\ell_1}}{\sqrt{s}}$

→ From part (5) & (7), we have,

$$\|h_{(T_0 \cup T_1)^c}\|_2 \le s^{-1/2}\left(\|h_{T_0}\|_{\ell_1} + 2\|x_{T_0^c}\|_{\ell_1}\right) \quad \text{—— (11)}$$

Now, for a s sparse vector v by cauchy schwartz inequality, we have

$$\|v\|_1 = \sum_{i=1}^{s} |v_i| = \sum_{i=1}^{s} |v_i| \cdot 1 \le \left(\sum_{i=1}^{s} |v_i|^2\right)^{1/2} \left(\sum_{i=1}^{s} 1^2\right)^{1/2}$$

$$\le \sqrt{s}\,\|v\|_2$$

∴ since h_{T_0} is s-sparse,

$$\|h_{T_0}\|_{\ell_1} \le s^{1/2}\|h_{T_0}\|_{\ell_2} \quad \text{—— (12)}$$

Using (12) & (11),

$$\|h_{(T_0 \cup T_1)^c}\|_2 \le s^{-1/2}\left(s^{1/2}\|h_{T_0}\|_{\ell_2} + 2\|x_{T_0^c}\|_{\ell_1}\right)$$

$$\le \|h_{T_0}\|_{\ell_2} + 2 s^{-1/2}\|x_{T_0^c}\|_{\ell_1}$$

$$\le \|h_{T_0}\|_{\ell_2} + 2e_0$$

Hence proved.

(9) Justify:

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle| \le \|\Phi h_{T_0 \cup T_1}\|_{\ell_2} \|\Phi h\|_{\ell_2} \le 2\epsilon\sqrt{1+\delta_{2s}}\,\|h_{T_0 \cup T_1}\|_{\ell_2}$$

→ By Cauchy-Schwartz inequality,

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle| \le \|\Phi h_{T_0 \cup T_1}\|_{\ell_2} \|\Phi h\|_{\ell_2} \quad \text{—— (13)}$$

$h = x^* - x$

PAGE NO.:
DATE: / /

From part (2), we have,

$$\|\phi h\|_{\ell_2} = \|\phi(x^* - x)\|_{\ell_2} \leq 2\epsilon \quad \text{—— (14)}$$

Now, $h_{T_0 \cup T_1}$ is a $2s$ sparse vector.
By RIP of order $2s$,

$$\|\phi h_{T_0 \cup T_1}\|_{\ell_2} \leq \sqrt{1+\delta_{2s}}\, \|h_{T_0 \cup T_1}\|_{\ell_2} \quad \text{—— (15)}$$

Combining (14), (15) & using in (13), we get,

$$|\langle \phi h_{T_0 \cup T_1}, \phi h \rangle| \leq 2\epsilon\sqrt{1+\delta_{2s}}\, \|h_{T_0 \cup T_1}\|_{\ell_2}$$

(10) Justify: $|\langle \phi h_{T_0}, \phi h_{T_j} \rangle| \leq \delta_{2s} \|h_{T_0}\|_{\ell_2} \|h_{T_j}\|_{\ell_2}$

→ Lemma 2.1 states
"for all x, x' supported on disjoint subsets $T, T' \subseteq \{1, 2, \ldots n\}$ with $|T| \leq s$, $|T'| \leq s'$,

$$|\langle \phi x, \phi x' \rangle| \leq \delta_{s+s'} \|x\|_{\ell_2} \|x'\|_{\ell_2} \text{"}$$

We use this lemma for $x = h_{T_0}$, $x' = h_{T_j}$
T_0 & T_j are disjoint for $j \neq 0$ & x, x' are both s-sparse.

$$\therefore |\langle \phi h_{T_0}, \phi h_{T_j} \rangle| \leq \delta_{2s} \|h_{T_0}\|_{\ell_2} \|h_{T_j}\|_{\ell_2}$$

(11) Justify: $\|h_{T_0}\|_{\ell_2} + \|h_{T_1}\|_{\ell_2} \leq \sqrt{2}\, \|h_{T_0 \cup T_1}\|_{\ell_2}$

→ T_0 & T_1 are disjoint. $h_{T_0 \cup T_1} = h_{T_0} + h_{T_1}$
By the QM-AM inequality, $\left\{ \frac{a+b}{2} \leq \sqrt{\frac{a^2+b^2}{2}} \right\}$

$$\frac{\|h_{T_0}\|_2 + \|h_{T_1}\|_2}{2} \leq \sqrt{\frac{\|h_{T_0}\|_{\ell_2}^2 + \|h_{T_1}\|_{\ell_2}^2}{2}} = \sqrt{\frac{\|h_{T_0 \cup T_1}\|_2^2}{2}}$$

$$\therefore \|h_{T_0}\|_2 + \|h_{T_1}\|_2 \leq \sqrt{2}\, \|h_{T_0 \cup T_1}\|_2$$

(12) Justify:

$$(1-\delta_{2S})\|h_{T_0\cup T_1}\|_{\ell_2}^2 \le \|\phi h_{T_0\cup T_1}\|_{\ell_2}^2 \le \|h_{T_0\cup T_1}\|_{\ell_2}\left(2\epsilon\sqrt{1+\delta_{2S}} + \sqrt{2}\delta_{2S}\sum_{j\ge 2}\|h_{T_j}\|_{\ell_2}\right)$$

$\rightarrow$ The first inequality $(1-\delta_{2S})\|h_{T_0\cup T_1}\|_{\ell_2}^2 \le \|\phi h_{T_0\cup T_1}\|_{\ell_2}^2$

is same as the RIP for order 2S, as $h_{T_0\cup T_1}$ is a 2S sparse vector.

From the paper, $\phi h_{T_0\cup T_1} = \phi h - \sum_{j\ge 2}\phi h_{T_j}$

$$\|\phi h_{T_0\cup T_1}\|_{\ell_2}^2 = \langle \phi h_{T_0\cup T_1}, \phi h\rangle - \langle \phi h_{T_0\cup T_1}, \sum_{j\ge 2}\phi h_{T_j}\rangle$$

Using inequality $(a-b \le |a|+|b|)$

$$\|\phi h_{T_0\cup T_1}\|_{\ell_2}^2 \le |\langle \phi h_{T_0\cup T_1}, \phi h\rangle| + |\langle \phi h_{T_0\cup T_1}, \sum_{j\ge 2}\phi h_{T_j}\rangle| \quad (16)$$

From part (9), we have

$$|\langle \phi h_{T_0\cup T_1}, \phi h\rangle| \le 2\epsilon\sqrt{1+\delta_{2S}}\,\|h_{T_0\cup T_1}\|_{\ell_2} \quad (17)$$

Now, $|\langle \phi h_{T_0\cup T_1}, \sum_{j\ge 2}\phi h_{T_j}\rangle| = \sum_{j\ge 2}|\langle \phi h_{T_0\cup T_1}, \phi h_{T_j}\rangle|$

T_0, T_1 are disjoint. From part (10), we have

$$|\langle \phi h_{T_0}, \phi h_{T_j}\rangle| \le \delta_{2s} \|h_{T_j}\|_{\ell_2} \|h_{T_0}\|_{\ell_2}$$

& $$|\langle \phi h_{T_1}, \phi h_{T_j}\rangle| \le \delta_{2s} \|h_{T_j}\|_{\ell_2} \|h_{T_1}\|_{\ell_2}$$

On adding above & noting $h_{T_0 \cup T_1} = h_{T_0} + h_{T_1}$

$$|\langle \phi h_{T_0 \cup T_1}, \phi h_{T_j}\rangle| \le \delta_{2s} \|h_{T_j}\|_{\ell_2} (\|h_{T_0}\|_{\ell_2} + \|h_{T_1}\|_{\ell_2})$$

Summing for $j \geq 2$.

$$\sum_{j \geq 2} |\langle \phi h_{T_0 \cup T_1}, \phi h_{T_j}\rangle| \le \sum_{j \geq 2} \delta_{2s} \|h_{T_j}\|_{\ell_2} (\|h_{T_0}\|_2 + \|h_{T_1}\|_2)$$

Now from part (11), $\|h_{T_0}\|_2 + \|h_{T_1}\|_2 \le \sqrt{2}\, \|h_{T_0 \cup T_1}\|_2$

$$\therefore \sum_{j \geq 2} |\langle \phi h_{T_0 \cup T_1}, \phi h_{T_j}\rangle| \le \sum_{j \geq 2} \delta_{2s} \|h_{T_j}\|_{\ell_2} (\sqrt{2}\, \|h_{T_0 \cup T_1}\|_{\ell_2}) \quad (18)$$

Adding (17) & (18) and using 16, we get,

$$\|\phi h_{T_0 \cup T_1}\|_2^2 \le \|h_{T_0 \cup T_1}\|_{\ell_2} \left(2\epsilon\sqrt{1+\delta_{2s}} + \sqrt{2}\,\delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_{\ell_2}\right)$$

Hence proved.

(13) Justify: $\|h_{T_0 \cup T_1}\|_{\ell_2} \le \alpha\epsilon + \rho s^{-1/2} \|h_{T_0^c}\|_{\ell_1}$

$$\alpha = \frac{2\sqrt{1+\delta_{2s}}}{(1-\delta_{2s})}, \quad \rho = \frac{\sqrt{2}\,\delta_{2s}}{(1-\delta_{2s})}$$

(13)

→ From previous part,

$$(1-\delta_{2S})\|h_{T_0 \cup T_1}\|_{\ell_2}^2 \leq \|h_{T_0\cup T_1}\|_{\ell_2}\left(2\epsilon\sqrt{1+\delta_{2S}} + \sqrt{2}\,\delta_{2S}\sum_{j\geq 2}\|h_{T_j}\|_{\ell_2}\right) \quad (19)$$

From part (4), we have,

$$\sum_{j\geq 2}\|h_{T_j}\|_{\ell_2} \leq S^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

Now, ~~(18) becomes~~, using above (19) becomes,

$$(1-\delta_{2S})\|h_{T_0\cup T_1}\|_{\ell_2}^2 \leq \|h_{T_0\cup T_1}\|_{\ell_2}\left(2\epsilon\sqrt{1+\delta_{2S}} + \sqrt{2}\,\delta_{2S}\, S^{-1/2}\|h_{T_0^c}\|_{\ell_1}\right)$$

On simplifying, & rearranging in proper form

$$\|h_{T_0\cup T_1}\|_{\ell_2} \leq \frac{2\epsilon\sqrt{1+\delta_{2S}}}{(1-\delta_{2S})} + \frac{\sqrt{2}\,\delta_{2S}}{(1-\delta_{2S})} S^{-1/2}\|h_{T_0^c}\|_{\ell_1}$$

Hence proved, $\|h_{T_0\cup T_1}\|_{\ell_2} \leq \alpha\epsilon + \rho S^{-1/2}\|h_{T_0^c}\|_{\ell_1}$,

where $\alpha = \dfrac{2\sqrt{1+\delta_{2S}}}{1-\delta_{2S}}$, $\rho = \dfrac{\sqrt{2}\,\delta_{2S}}{1-\delta_{2S}}$

(14) Justify: $\|h_{T_0\cup T_1}\|_{\ell_2} \leq \alpha\epsilon + \rho\|h_{T_0\cup T_1}\|_{\ell_2} + 2\rho e_0$

→ From part 7 {eqn (12) in paper}, i.e.

$$\|h_{T_0^c}\|_{\ell_1} \leq \|h_{T_0}\|_{\ell_1} + 2\|x_{T_0^c}\|_{\ell_1} = \|h_{T_0}\|_{\ell_1} + 2S^{1/2}e_0$$

Using this, in previous part (13), we get,

$$\|h_{T_0\cup T_1}\|_{\ell_2} \leq \alpha\epsilon + \rho S^{-1/2}\left(\|h_{T_0}\|_{\ell_1} + 2S^{1/2}e_0\right)$$

$$\leq \alpha\epsilon + \rho S^{-1/2}\|h_{T_0}\|_{\ell_1} + 2\rho e_0$$

From part (8), equ^n 12,

$$\|h_{T_0}\|_{\ell_1} \le s^{1/2} \|h_{T_0}\|_{\ell_2}$$

Using this, we get,

$$\|h_{T_0 \cup T_1}\|_{\ell_2} \le \alpha \epsilon + \rho \|h_{T_0}\|_{\ell_2} + 2\rho e_0$$

Now, using $\|h_{T_0}\|_{\ell_2} \le \|h_{T_0 \cup T_1}\|_{\ell_2}$, we get,

$$\|h_{T_0 \cup T_1}\|_{\ell_2} \le \alpha \epsilon + \rho \|h_{T_0 \cup T_1}\|_{\ell_2} + 2\rho e_0$$

Hence proved

(15) Justify:

$$\|h\|_{\ell_2} \le \|h_{T_0 \cup T_1}\|_{\ell_2} + \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} \le 2\|h_{T_0 \cup T_1}\|_{\ell_2} + 2e_0$$

$$\le 2(1-\rho)^{-1}(\alpha \epsilon + (1+\rho)e_0)$$

(15)

→ We have $h = h_{T_0 \cup T_1} + h_{(T_0 \cup T_1)^c}$

By triangle inequality,

$$\|h\|_{\ell_2} \leq \|h_{T_0 \cup T_1}\|_{\ell_2} + \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} \quad \text{— } (20)$$

From part (8) we have,

$$\|h_{(T_0 \cup T_1)^c}\|_{\ell_2} \leq \|h_{T_0}\|_{\ell_2} + 2e_0$$

$$\therefore \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} \leq \|h_{T_0 \cup T_1}\|_{\ell_2} + 2e_0 \quad \{\text{since } \|h_{T_0}\|_2 \leq \|h_{T_0 \cup T_1}\|_2\}$$

Using above in ~~20 eq~~ (20),

$$\|h\|_{\ell_2} \leq 2\|h_{T_0 \cup T_1}\|_{\ell_2} + 2e_0 \quad \text{— } (21)$$

From part (14), we have,

$$\|h_{T_0 \cup T_1}\|_{\ell_2} \leq \alpha \epsilon + \rho \|h_{T_0 \cup T_1}\|_{\ell_2} + 2\rho e_0$$

$$\therefore \|h_{T_0 \cup T_1}\|_{\ell_2} \leq (1-\rho)^{-1}\{\alpha \epsilon + 2\rho e_0\}$$

Using above in (21),

$$\|h\|_{\ell_2} \leq 2(1-\rho)^{-1}(\alpha \epsilon + 2\rho e_0) + 2e_0$$

$$= 2(1-\rho)^{-1}(\alpha \epsilon + (1+\rho)e_0)$$

Hence all inequalities are justified.

(16) Justify:

$$\|h\|_{\ell_1} = \|h_{T_0}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} \leq 2(1+\rho)(1-\rho)^{-1} \|x_{T_0^c}\|_{\ell_1}$$

$\rightarrow$ From Lemma 2.2, we have,

$$\|h_{T_0}\|_{\ell_1} \leq \rho \|h_{T_0^c}\|_{\ell_1}$$

$$\therefore \|h\|_{\ell_1} = \|h_{T_0}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} \leq (1+\rho)\|h_{T_0^c}\|_{\ell_1} \quad (22)$$

Now, ~~using~~ from the paper we have the result,

$$\|h_{T_0^c}\|_{\ell_1} \leq 2(1-\rho)^{-1} \|x_{T_0^c}\|_{\ell_1}$$

Using this in (22), we get,

$$\|h\|_{\ell_1} \leq 2(1+\rho)(1-\rho)^{-1} \|x_{T_0^c}\|_{\ell_1}$$

Hence proved.

Question 2

Consider compressive measurements

$$y = \Phi x + \eta$$

of a purely sparse signal x, where $\|\eta\|_2 \leq \varepsilon$.

a) Let $\tilde{x}$ be the oracular solution, i.e. $\tilde{x}$ is the solution that we could obtain if we knew in advance the indices (set S) of the non-zero elements of x.

From the definition of $\tilde{x}$ we have

$$y = \Phi_S \tilde{x} \quad \longrightarrow (i)$$

... (Φ_S is the submatrix formed by only considering the columns corresponding to indices in set S)

Since we know that $\Phi_S^T \Phi_S$ is invertible, we can directly write the solution for $\tilde{x}$ using pseudo inverse

$$\therefore \quad \tilde{x} = (\Phi_S^T \Phi_S)^{-1} \Phi_S^T y$$

... (Known solution for linear regression)

i.e. $\underline{\underline{\tilde{x} = \Phi_S^{+} y}}$ ... $(\Phi_S^{+} = (\Phi_S^T \Phi_S)^{-1} \Phi_S^T)$

b) $\|\tilde{x} - x\|_2 = \|\Phi_S^{+} y - x\|_2$

$$= \|\Phi_S^{+}(\Phi x + \eta) - x\|_2$$

Now, since x is purely sparse with at most $|S|$ non-zero elements, we can write

$$\Phi x = \Phi_S x \quad \text{... (By definition of } \Phi_S)$$

$$\therefore \quad \|\tilde{x} - x\|_2 = \|\phi_S^{+}(\phi_S x + \eta) - x\|_2$$
$$= \|(\phi_s^T \phi_s)^{-1} \phi_s^T \phi x + \phi_s^{+} \eta - x\|_2$$
$$= \|x + \phi_s^{+} \eta - x\|_2$$
$$= \|\phi_s^{+} \eta\|$$

Now, by definition of induced matrix norms

$$\|\tilde{x} - x\|_2 = \|\phi_s^{+} \eta\|_2 \leq \|\phi_s^{+}\|_2 \|\eta\|_2 \quad \text{——— (ii)}$$

$\ldots$ (where $\|\phi_s^{+}\|_2$ is the matrix 2-norm of ϕ_s^{+} which is defined as

$$\|\phi_s^{+}\|_2 = \max_{x \neq 0} \frac{\|\phi_s^{+} x\|_2}{\|x_2\|}$$)

From Linear algebra, we know that the maximum amplification caused by a matrix along some basis vector is given by the largest singular value and it follows

$$\sigma_{1, \phi_s^{+}} = \|\phi_s^{+}\|_2 \longleftarrow \text{induced 2 norm of } \phi_s^{+}$$

largest singular value of ϕ_s^{+} ——— (iii)

In the question, the largest singular value of matrix X is denoted by $\|X\|_2$. However, since the induced 2 norm of the matrix is same as the largest singular value, we have

$$\|\tilde{x} - x\|_2 \leq \|\phi_s^{+}\|_2 \|\eta\|_2 \quad \ldots \text{(Using (ii) and (iii))}$$

——— (iv)

c.) $K = |S|$

i.e. x is a k-sparse vector

$\Rightarrow$ $\tilde{x} - x$ can be atmost $2k$-sparse vector

We are given that δ_{2K} is the RIC of Φ of order $2K$

$$\Rightarrow (1-\delta_{2K})\|\tilde{x}-x\|_2^2 \leq \|\Phi(\tilde{x}-x)\|_2^2 \leq (1+\delta_{2K})\|\tilde{x}-x\|_2^2 \quad \text{———} (v)$$

Now, $\|\Phi(\tilde{x}-x)\|_2^2 = \|\Phi\tilde{x} - \Phi x\|_2^2$

$$= \|y - (y - \eta)\|_2^2$$

$$= \|\eta\|_2^2 \quad \text{———} (vi)$$

Using one of the inequality from (v) with (vi)

$$\|\eta\|_2^2 \leq (1+\delta_{2K})\|\tilde{x}-x\|_2^2$$

$$\Rightarrow \frac{\|\eta\|_2^2}{1+\delta_{2K}} \leq \|\tilde{x}-x\|_2^2$$

$$\Rightarrow \frac{\|\eta\|_2^2}{1+\delta_{2K}} \leq \|\Phi_S^+\|_2^2 \|\eta\|_2^2 \quad \text{--- (from (iv))}$$

$$\Rightarrow \|\Phi_S^+\|_2 \geq \frac{1}{\sqrt{1+\delta_{2K}}} \quad \text{———} (vii)$$

Similarly using other inequality in (v) we get

$$\|\Phi_S^+\|_2 \leq \frac{1}{\sqrt{1-\delta_{2K}}} \quad \text{———} (viii)$$

Using (vii) and (viii) we get

$$\frac{1}{\sqrt{1+\delta_{2K}}} \leq \|\Phi_S^{+}\|_2 \leq \frac{1}{\sqrt{1-\delta_{2K}}} \quad \cdots \left(\begin{array}{c}\text{Here } \|\Phi_S^{+}\|_2 \text{ is} \\ \text{the largest} \\ \text{singular value} \\ \text{of } \Phi_S^{+} \text{ as} \\ \text{mentioned earlier}\end{array}\right)$$

d.) Using the above result with (iv)

$$\frac{\|\eta\|_2}{\sqrt{1+\delta_{2K}}} \leq \|x-\tilde{x}\|_2 \leq \frac{\|\eta\|_2}{\sqrt{1-\delta_{2K}}}$$

$$\Rightarrow \frac{\varepsilon}{\sqrt{1+\delta_{2K}}} \leq \|x-\tilde{x}\|_2 \leq \frac{\varepsilon}{\sqrt{1-\delta_{2K}}} \quad \cdots \left(\text{Given } \|\eta\|_2^2 \leq \varepsilon\right) \longrightarrow (ix)$$

Also, from Theorem 3 we know that

$$\|x-x^*\|_2 \leq C_0 K^{-1/2} \|x-\tilde{x}\|_1 + C_1 \varepsilon \longrightarrow (x)$$

Also, for a 2K sparse vector we have the following relation between 1 and 2 norm

$$\|x-\tilde{x}\|_1 \leq \sqrt{2K}\, \|x-\tilde{x}\|_2 \longrightarrow (xi)$$

Combining (x) and (xi) we get

$$\|x-x^*\|_2 \leq \sqrt{2}\, C_0 \|x-\tilde{x}\|_2 + C_1 \varepsilon$$

Using (ix) we get

$$\|x - x^*\|_2 \leq \sqrt{2}\, C_0 \frac{\varepsilon}{\sqrt{1-\delta_{2k}}} + C_1 \varepsilon$$

$$\Rightarrow \quad \|x - x^*\|_2 \leq \left[\left(\frac{\sqrt{2}\, C_0}{\sqrt{1-\delta_{2k}}} + C_1\right)\sqrt{1+\delta_{2k}}\right] \|x - \tilde{x}\|_2$$

... (Using (ix))

Hence, we can see that

$$\|x - x^*\|_2 \leq C \, \|x - \tilde{x}\|_2$$

where $C = \left(\frac{\sqrt{2}\, C_0}{\sqrt{1-\delta_{2k}}} + C_1\right)\sqrt{1+\delta_{2k}}$

Hence proved.

CS 754 HW 2

Q3)

Soln – Given: $s < t$; s and t are positive integers
δ_s & δ_t are RIC of order s & t respectively.
Let x_s and x_t be s-sparse & t-sparse vectors.
By the RIP of order t,

$$(1-\delta_t)\|x_t\|_2^2 \leq \|\Phi x_t\|_2^2 \leq (1+\delta_t)\|x_t\|_2^2 \quad \text{— (1)}$$

Since $s < t$, any s-sparse vector is also t-sparse vector. Thus, x_s satisfies (1).

$$(1-\delta_t)\|x_s\|_2^2 \leq \|\Phi x_s\|_2^2 \leq (1+\delta_t)\|x_s\|_2^2 \quad \text{— (2)}$$

But from the definition of RIC of order s,
δ_s is the smallest δ which satisfies,

$$(1-\delta)\|x_s\|_2^2 \leq \|\Phi x_s\|_2^2 \leq (1+\delta)\|x_s\|_2^2 \quad \text{— (3)}.$$

$\therefore$ From (2), δ_t also satisfies (3).
Hence, $\delta_s \leq \delta_t$.
It is proved that if $s < t$, $\delta_s \leq \delta_t$.

Question 4

a) Title : Sensing Matrix Design for Compressive Spectral Imaging via Binary Principal Component Analysis

Author List : Jonathan Monsalve, Hoover Rueda-Chacon, Henry Arguello

Published in : IEEE Transactions on Image Processing (Volume: 29)

Pages : 4003-4012

Date of publication: 19 December 2019

Link to paper : https://ieeexplore.ieee.org/document/8937037

b) Imaging System in the paper

Core idea of the paper:

The sensing matrices used in many compressive spectral imaging architectures are binary matrices whose elements can either be random or designed. However, some characteristics such as l_2-norm or second moment statistics can be lost when the dimensionality reduction is performed. We also know that PCA is a dimensionality reduction technique that minimizes the least squared error between spectral data and its low-dimensional projection. Hence PCA can be used to guide the compressed sensing acquisition process by designing the binary sensing matrix.

First, a set of compressive measurements obtained with random sensing matrices is used to rapidly estimate the covariance matrix associated with the spectral data. Then, a new sensing matrix is designed by solving a non-convex optimisation problem that finds a set of binary vectors that approximate the principal components of the covariance matrix, thus maximising the expectation of data variance

Imaging System used

The imaging systems that can be used to implement this approach are 3D-CASSI and the DD-CASSI. (CASSI stands for coded aperture snapshot spectral imager) The complete flowchart for sensing and reconstruction is as shown below.

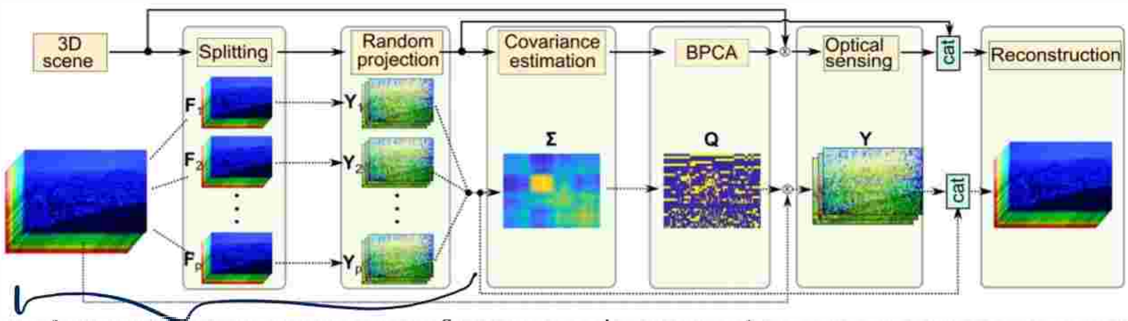
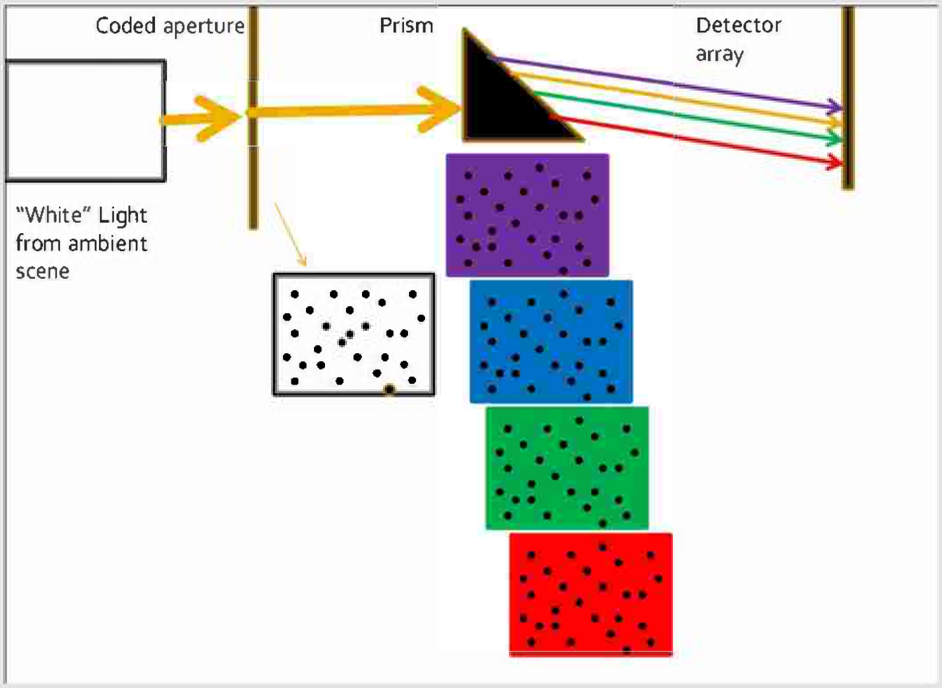


Fig. 2. Flowchart for the sensing and reconstruction algorithm. Solid lines show a block diagram of the procedure and dotted lines show graphically the same procedure. First, the image is divided into p subsets and projected using random matrices. The covariance matrix is estimated using these random projections and it is used to design the binary sensing matrix.

Imaging System

This is similar to the CASSI system that has been covered in the lectures



Assume we want to measure a hyperspectral data cube given as $X \in \mathbb{R}^{N_x \times N_y \times N_\lambda}$ where data at each wavelength is a 2D image of size $N_x \times N_y$ and number of wavelengths is N_λ.

In a CASSI camera, each image X_j, $1 \leq j \leq N_\lambda$, is multiplied by the binary code given as $C \in \{0,1\}^{N_x \times N_y}$ yielding an image

$$\hat{X}_j = X_j \cdot C$$

Hence, as derived in lectures, the measurement by the CASSI system in a single 2D snapshot is given as follows (superposition of coded data from all wavelengths)

$$M(x,y) = \sum_{j=1}^{N_\lambda} X_j(x - l_j, y) \cdot C(x - l_j, y)$$

... (Exact formulation is in the lecture slides)

c.) Mathematical expression for matrix quality measure being optimized in the paper, along with various constraints on the matrix

Solution.) The columns q_j of the sensing matrix Q are obtained by solving the optimization problem

$$q_j = \underset{q_j, b_j}{\operatorname{argmax}} \; q_j^T \Sigma q_j$$

$$\text{subject to} \quad q_j^k \in \{0, \frac{1}{\sqrt{b_j}}\}$$

where q_j^k is the k^{th} entry of the j^{th} column of Q and Σ is the covariance matrix

The above problem aims to estimate the subspace spanned by the binary vectors q_j, which maximize the variance of data concentrated in the matrix

Note: The covariance matrix Σ is unknown, since it depends on data. Thus, a set of random projections must be acquired at first in order to estimate it. This paper uses CPPCA ("Compressive-projection principal component analysis") approach given in a reference paper.

d.) Mention the optimization technique.

Solution)

The above mentioned optimization problem is approximately solved using a greedy-search-based approach since it iteratively looks for the best position with the vector that maximizes the objective function.

The complete algorithm is as follows:

Algorithm 1 B-PCA: Binary PCA estimation

input: $\mathbf{\Sigma}_1, \tilde{m}$
for $j = 1$ to $\tilde{m}$ **do**
  $\mathbf{q}_j \leftarrow 0$; $max \leftarrow 0$; $list = 1, 2, \cdots, l$
  for $k = 1$ to round$(l/\tilde{m})$ **do** *impose transmit.* $1/\tilde{m}$
    for $i =$ each element in $list$ **do**
      $\mathbf{q}_j^i \leftarrow 1$ *place a one in the* i^{th} *position*
      $c^k \leftarrow \frac{\mathbf{q}_j^T}{\|\mathbf{q}_j\|}\mathbf{\Sigma}_j\frac{\mathbf{q}_j}{\|\mathbf{q}_j\|}$ *objective function* (7)
      if $c^k > max$ **then**
        $max = c^k$; $index = i$
      end if
      $\mathbf{q}_j^i \leftarrow 0$ *remove the one value*
    end for
    $\mathbf{q}_j^{index} \leftarrow 1$ *place the one in the best position*
    $list.remove(index)$; $index \leftarrow 0$
  end for
  $\mathbf{P} \leftarrow \frac{\mathbf{q}_j\mathbf{q}_j^T}{\|\mathbf{q}_j\|^2}$
  $\mathbf{\Sigma}_{j+1} \leftarrow \mathbf{\Sigma}_j - \mathbf{\Sigma}_j\mathbf{P} - \mathbf{P}\mathbf{\Sigma}_j + \mathbf{P}\mathbf{\Sigma}_j\mathbf{P}$
end for
output: $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \cdots \mathbf{q}_{\tilde{m}}]$

e) Briefly describe the improvements due to this design as compared to a random design.

Solution)

The reconstruction performance of this approach was tested on two datasets (Urban Image and Pavia Image)

The comparative plot for average PSNR of the reconstructed hyperspectral datasets is shown below.

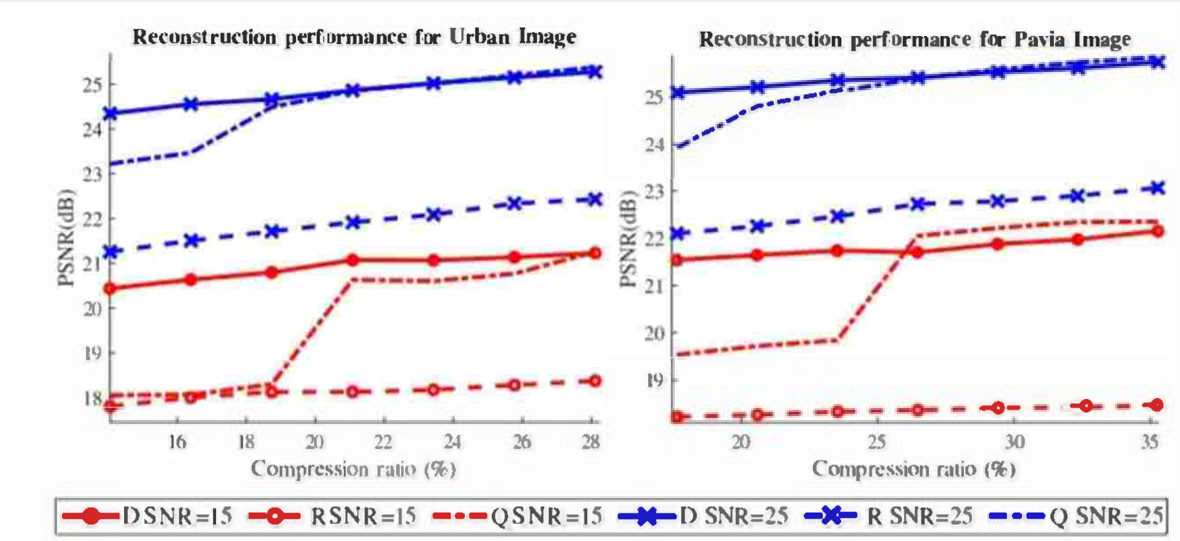


Fig. 8. Average PSNR of the reconstructed hyperspectral datasets for various compression ratios, at 2 noise scenarios with SNR = 15 and 25 dB. "D" stands for designed (solid lines), "Q" for those proposed in [18] (Qpca) (dot-dashed lines) and "R" for random (dotted lines) matrices. Different colors represent different noise levels. (Left) Urban dataset. (Right) Pavia dataset

From the above plots we can easily see that the performance of the designed matrices achieves an overall 3 dB improvement in the reconstruction quality compared with conventional random sensing matrices for the same compression ratio on both the hyperspectral datasets.

Q5)

Soln.

P1 : $\min\limits_{x} \|x\|_1$ s.t. $\|y-\Phi x\|_2 \leq \epsilon$

LASSO : $\min\limits_{x} \left(\|y-\Phi x\|_2^2 + \lambda \|x\|_1 \right)$

Statement: If x is minimizer of LASSO for some $\lambda > 0$, there exists some ϵ for which x is also minimizer of P1.

Proof: x is minimizer of LASSO problem for some $\lambda > 0$.

Let $\epsilon' = \|y - \Phi x\|_2$

Consider $z \in \mathbb{R}^n$ such that $\|y-\Phi z\|_2 \leq \epsilon'$

For $\epsilon = \epsilon'$, z satisfies constraint of P1.

Now, since x is minimizer of LASSO,

$$\|y-\Phi x\|_2^2 + \lambda\|x\|_1 \leq \|y-\Phi z\|_2^2 + \lambda \|z\|_1 \quad \text{①}$$

We have $\|y-\Phi z\|_2 \leq \epsilon' = \|y-\Phi x\|_2$

$$\therefore \|y-\Phi z\|_2^2 \leq \|y-\Phi x\|_2^2$$

~~From ①~~, $\therefore \|y-\Phi z\|_2^2 + \lambda\|z\|_1 \leq \|y-\Phi x\|_2^2 + \lambda\|z\|_1 \quad \text{②}$

From ① & ②,

$$\|y-\Phi x\|_2^2 + \lambda\|x\|_1 \leq \|y-\Phi x\|_2^2 + \lambda\|z\|_1$$

$$\therefore \|x\|_1 \leq \|z\|_1$$

Now since x also satisfies constraint of P1 for ϵ', and $\|x\|_1 \leq \|z\|_1$ for all $z \in \mathbb{R}^n$ satisfying the constraint, x is also minimizer of P1.

Hence, proved.

Question 6

Given: There are n subjects being tested by Dorfman pooling and only $K << n$ out of these are infected.
In first round the n subjects are divided into groups of size g each
Hence, we have $\frac{n}{g}$ groups of size g each

a.) Find average number of tests required to perform Dorfman pooling.

Solution)
In the first round we need to perform test on all the $\frac{n}{g}$ groups which are of size g each.

After the first round a given group can either be positive (i.e. atleast one member of the group is positive) or it can be negative (i.e. no member in the group is positive)

- If a group is positive all its members are tested.
Hence for every group which is positive we need total $1 + g$ tests.
(Here 1 corresponds to the complete group test in first round and g tests for each of the members in the second round)

Next, we calculate the probability that a given group is positive.

$$P_{+ve\ group} = \mathbb{P}(\text{Atleast one member in the group is positive})$$

$$= 1 - \mathbb{P}(\text{none of the members in the group is positive})$$

$$= 1 - \left(1 - \frac{k}{n}\right)^g$$

... (Since all the n members are equally likely to be infected independent of each other)

- If a group is tested negative, we do not need to test any member of the group. Hence for a negative group, we need just 1 test in the first round

The probability that a group is negative is

$$\mathbb{P}_{\substack{-ve\\group}} = 1 - \mathbb{P}_{\substack{+ve\\group}}$$

$$= 1 - \left(1 - \left(1 - \frac{k}{n}\right)^g\right)$$

$$= \left(1 - \frac{k}{n}\right)^g$$

Hence, average number of tests for a single group

$$= \begin{pmatrix}\text{No. of tests}\\ \text{if group}\\ \text{is +ve}\end{pmatrix} \times \mathbb{P}_{\substack{+ve\\group}} + \begin{pmatrix}\text{No. of tests}\\ \text{if group}\\ \text{is -ve}\end{pmatrix} \times \mathbb{P}_{\substack{-ve\\group}}$$

$$= (1+g)\left(1 - \left(1 - \frac{k}{n}\right)^g\right) + \left(1 - \frac{k}{n}\right)^g$$

$$= 1 - \cancel{\left(1 - \frac{k}{n}\right)^g} + g\left(1 - \left(1 - \frac{k}{n}\right)^g\right) + \cancel{\left(1 - \frac{k}{n}\right)^g}$$

$$= 1 + g\left(1 - \left(1 - \frac{k}{n}\right)^g\right)$$

Hence, we can find average number of tests required as

$$\text{Avg. number of tests} = \left(\text{No. of groups}\right) \times \left(\text{Avg. no. of tests per group}\right)$$

$$= \frac{n}{g} \times \left(1 + g\left(1 - \left(1 - \frac{k}{n}\right)^{g}\right)\right)$$

$$= \underline{\frac{n}{g} + n\left(1 - \left(1 - \frac{k}{n}\right)^{g}\right)}$$

b.) What is the worst case?

Assumption: $K < \frac{n}{g}$

i.e. the number of infected individuals are less than the number of groups

Solution) The worst case will happen, i.e. we will need maximum number of tests when all the k infected individuals will be present in different groups.

This is because in such a case, all members of a group which has an infected member will be tested in the second round of Dorfman pooling. Hence, the total number of tests will be more compared to a situation where any of the group will have more than one infected individuals.

c.) What is the optimal group size in the worst case?

Solution.)

In the worst case, we will have to conduct $\left(\frac{n}{g}\right)$ tests corresponding to the number of groups in the first round of Dorfman pooling.

In the second round, since we are assuming that all the K infected individuals are present is different groups (in worst case), we will have to conduct

$K \times g$ test (i.e. all the k positive groups should be tested)

$\therefore$ Total number of tests in worst case,

$$N = \frac{n}{g} + kg$$

Since we want to minimize N w.r.t g to find optimal group size g^*,

$$\frac{\partial N}{\partial g} = \frac{\partial}{\partial g}\left(\frac{n}{g} + kg\right) = 0$$

$$\Rightarrow -\frac{n}{g^2} + K = 0$$

$$\Rightarrow \underline{\underline{g^* = \sqrt{\frac{n}{K}}}}$$

Hence, the optimal group size g^* in worst case is $\underline{g^* = \sqrt{\frac{n}{K}}}$