Optimization Theory HW2

RE6124019 吳明軒

November 2023

1. Given the Hand: [2, 2, 2, 3, 4, 5, 6, 7, 7, 7, 8, 8, 8, 9, 9, 9]
Prove that the Winning tiles are given by:[1, 2, 3, 4, 5, 6, 7, 8, 9]

1: [{2, 2}, {1, 2, 3}, {4, 5, 6}, {7, 7, 7}, {8, 8, 8}, {9, 9, 9}]
2: [{2, 2, 2}, {2, 3, 4}, {5, 6, 7}, {7, 7}, {8, 8, 8}, {9, 9, 9}]
3: [{2, 2, 2}, {3, 3}, {4, 5, 6}, {7, 7, 7}, {8, 8, 8}, {9, 9, 9}]
4: [{2, 2}, {2, 3, 4}, {4, 5, 6}, {7, 7, 7}, {8, 8, 8}, {9, 9, 9}]
5: [{2, 2, 2}, {3, 4, 5}, {5, 6, 7}, {7, 7}, {8, 8, 8}, {9, 9, 9}]
6: [{2, 2, 2}, {3, 4, 5}, {6, 6}, {7, 7, 7}, {8, 8, 8}, {9, 9, 9}]
7: [{2, 2}, {2, 3, 4}, {5, 6, 7}, {7, 7, 7}, {8, 8, 8}, {9, 9, 9}]
8: [{2, 2, 2}, {3, 4, 5}, {6, 7, 8}, {7, 7}, {8, 8, 8}, {9, 9, 9}]
9: [{2, 2, 2}, {3, 4, 5}, {6, 7, 8}, {7, 8, 9}, {7, 8, 9}, {9, 9}]

2. In Hong Kong(x = 13), an elegant and efficient RH is
(1, 1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 9, 9)
Amazingly, the winning tiles are:(1, 2, 3, 4, 5, 6, 7, 8, 9)!

1: $[\{1,1,1\},\{1,2,3\},\{4,5,6\},\{7,8,9\},\{9,9\}]$
2: $[\{1,1,1\},\{2,2\},\{3,4,5\},\{6,7,8\},\{9,9,9\}]$
3: $[\{1,1\},\{1,2,3\},\{3,4,5\},\{6,7,8\},\{9,9,9\}]$
4: $[\{1,1,1\},\{2,3,4\},\{4,5,6\},\{7,8,9\},\{9,9\}]$
5: $[\{1,1,1\},\{2,3,4\},\{5,5\},\{6,7,8\},\{9,9,9\}]$
6: $[\{1,1\},\{1,2,3\},\{4,5,6\},\{6,7,8\},\{9,9,9\}]$
7: $[\{1,1,1\},\{2,3,4\},\{5,6,7\},\{7,8,9\},\{9,9\}]$
8: $[\{1,1,1\},\{2,3,4\},\{5,6,7\},\{8,8\},\{9,9,9\}]$
9: $[\{1,1\},\{1,2,3\},\{4,5,6\},\{7,8,9\},\{9,9,9\}]$

3.

1. Please name an advantage of DL;
Won't have to deal with relatively complicated mathematics.

2. Please name an advantage of CO;
Won't have to rely on big data.

3. Please name a disadvantage of DL;
Will have to rely on big data.

4. Please name a disadvantage of CO.
Will have to deal with relatively complicated mathematics.

5. How the ADMM-Adam theory avoids using big data?
Training deep learning models with small datasets may lead to poor results, but even in such poor outcomes, there is still valuable information. The ADMM-Adam theory employs the design of Q quadratic norm to extract this crucial information.

6. How the ADMM-Adam theory avoids using heavy math?
The ADMM-Adam theory designs a very simple mathematical modulator. Compared to conventional TV, the non-differentiable modulator or self-similarity, which includes a graph modulator, is a straightforward function. It avoids the use of complex mathematics.

4. Prove the equality $\|x - x_{DL}\|_Q^2 = \|S - S_{DL}\|_F^2$

[Hint: E is a semiunitary matrix, i.e., $E^T E = I_N$.]

Let $X \in \mathbb{R}^{ML}$ be an M-band hyperspectral image with L pixels.

Let $E \in \mathbb{R}^{ML}$ have columns with orthonormal basis, representing the N-dimensional hyperspectral subspace.

The hyperspectral data is given by $X = ES$ with $S \in \mathbb{R}^{NL}$.

Similarly, for the denoised version X_{DL}, we have $X_{DL} = ES_{DL}$ with $S_{DL} \in \mathbb{R}^{NL}$.

Choosing the PSD matrix $Q = I_L \otimes (EE^T)$ simplifies the analysis.

Proof :

$$
\begin{aligned}
&\|\mathrm{x} - \mathrm{x}_{DL}\|_Q^2 = \|\mathrm{vec}(X) - \mathrm{vec}(X_{DL})\|_Q^2 \\
&= \|\mathrm{vec}(X - X_{DL})\|_Q^2 \\
&= \mathrm{vec}(X - X_{DL})^T Q \mathrm{vec}(X - X_{DL}) \\
&= \mathrm{vec}(X - X_{DL})^T (I_L \otimes (EE^T)) \mathrm{vec}(X - X_{DL}) \quad \because Q = I_L \otimes (EE^T) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(EE^T (X - X_{DL})) I_L \quad \because (B^T \otimes A)(X) = \mathrm{vec}(A^T B) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(E(S - S_{DL})) I_L \quad \because X - X_{DL} = ES - ES_{DL} = E(S - S_{DL}) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(E) \mathrm{vec}(S - S_{DL}) I_L \quad \because \mathrm{vec}(A^T B) = (\mathrm{vec}(B)^T \otimes I_m) \mathrm{vec}(A) \\
&= \mathrm{vec}(X - X_{DL})^T (E^T \otimes I_L) \mathrm{vec}(S - S_{DL}) \quad \because \mathrm{vec}(E) = E^T \otimes I_L \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}((E^T \otimes I_L)(S - S_{DL})) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}((E^T \otimes I_L)(E(S - S_{DL}))) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(E^T E(S - S_{DL})) \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(I_N (S - S_{DL})) \quad \because E^T E = I_N \\
&= \mathrm{vec}(X - X_{DL})^T \mathrm{vec}(S - S_{DL}) \\
&= \|\mathrm{vec}(S - S_{DL})\|_F^2
\end{aligned}
$$