最佳化 HW3 徐仁瓏 RE6121011

- (Homework)
  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]

Proof:

— 1,2,3	— 2,2,2	— 2,2,2	— 2,3,4	— 2,2,2
2,2	2,3,4	3,3	2,2	3,4,5
4,5,6	5,6,7	4,5,6	4,5,6	5,6,7
7,7,7	7,7	7,7,7	7,7,7	7,7
8,8,8	8,8,8	8,8,8	8,8,8	8,8,8
9,9,9	9,9,9	9,9,9	9,9,9	9,9,9

— 2,2,2	— 2,2	— 2,2,2	— 2,2,2
3,4,5	2,3,4	3,4,5	3,4,5
6,6	5,6,7	6,7,8	6,7,8
7,7,7	7,7,7	7,7	7,8,9
8,8,8	8,8,8	8,8,8	7,8,9
9,9,9	9,9,9	9,9,9	9,9

- (Homework)
  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)!
  Prove it.

— 1,1,1	— 1,1,1	— 1,1	— 1,1,1	— 1,1,1
1,2,3	2,2	1,2,3	2,3,4	2,3,4
4,5,6	3,4,5	3,4,5	4,5,6	5,5
7,8,9	6,7,8	6,7,8	7,8,9	6,7,8
9,9	9,9,9	9,9,9	9,9	9,9,9

— 1,1	— 1,1,1	— 1,1,1	— 1,1
1,2,3	2,3,4	2,3,4	1,2,3
4,5,6	5,6,7	5,6,7	4,5,6
6,7,8	7,8,9	8,8	7,8,9
9,9,9	9,9	9,9,9	9,9,9

- ADMM and Adam are two optimizers of paramount importance in convex optimization (CO) and deep learning (DL), respectively.

  (Homework) Please check this YouTube video:

  https://www.youtube.com/watch?v=1isVbbMsGs4&ab_channel=YSFTaiwan

  and answer the following questions:

  1. please name an advantage of DL;
  2. please name an advantage of CO;
  3. please name a disadvantage of DL;
  4. please name a disadvantage of CO.
  5. how the ADMM-Adam theory avoids using big data?
  6. how the ADMM-Adam theory avoids using heavy math?

1、無艱深數学

2、只需单一數據

3、需大數據

4、需艱深數学

5、雖然小數據会使 DL 的 solution 非常爛，
但我们相信这麼爛的 solution 中还是包含著非常重要的 information，
所以我們可以透过 Q quadratic norm 將这重要的 information 萃取出來。

6、透过简单的 convex Q quadratic norm regularizer 可以避免使用
像 Graph-embedded SS 这种複雜的 regularizer，
儘管这个數学非常简单，还是成功地將这个理論应用來破解
一个具挑戰性的高維度的高光譜影像还原问題。

- (Homework) Prove the equality $\|\boldsymbol{x} - \boldsymbol{x}_{\mathrm{DL}}\|_{\boldsymbol{Q}}^2 = \|\boldsymbol{S} - \boldsymbol{S}_{\mathrm{DL}}\|_F^2$.

  [Hint: $\boldsymbol{E}$ is a semiunitary matrix, i.e., $\boldsymbol{E}^T\boldsymbol{E} = \boldsymbol{I}_N$.]

Let $x = \mathrm{vec}(X) \in R^{ML}$, where $X \in R^{M\times L}$ is an M-band hyperspectral image with L pixels.

Let the columns of $E \in R^{M\times N}$ represent an orthonormal basis of the N-dimensional hyperspectral subspace.

Then, the hyperspectral data can be represented as $X = ES$ for some coefficient matrix $S \in R^{N\times L}$

Similarly, we have $X_{DL} = ES_{DL}$ for some coefficient matrix $S_{DL} \in R^{N\times L}$

And by choosing the PSD matrix as $Q = I_L \otimes (EE^T)$,

we can prove :

$$\begin{aligned}
\| x - x_{DL} \|_Q^2 &= \| \mathrm{vec}(X) - \mathrm{vec}(X_{DL}) \|_Q^2 \\
&= \| \mathrm{vec}(X - X_{DL}) \|_Q^2 \qquad \because \text{攤開後相減} = \text{相減後攤開} \\
&= \mathrm{vec}(X - X_{DL})^T Q \, \mathrm{vec}(X - X_{DL}) \\
&= \mathrm{vec}(X - X_{DL})^T I_L \otimes (EE^T) \, \mathrm{vec}(X - X_{DL}) \qquad \because Q = I_L \otimes (EE^T)
\end{aligned}$$

$\because (B^T \otimes A)\,\mathrm{vec}(X) = \mathrm{vec}(AXB)$

by lemma Vectorization & Kronecker Product

$$= \mathrm{vec}(\underbrace{X - X_{DL}}_{M\times L})^T \, \mathrm{vec}(\underbrace{E}_{M\times N}\overbrace{E^T}^{N\times M}\underbrace{(X - X_{DL})}_{M\times L}\underbrace{I_L}_{L\times L \,\rightarrow\, M\times L})$$

$\because \mathrm{trace}(A'B) = \mathrm{vec}(A)'\mathrm{vec}(B)$ if A & B have same size.

$$\begin{aligned}
&= \mathrm{trace}\left[(X - X_{DL})^T \cdot EE^T (X - X_{DL}) I_L \right] \\
&= \mathrm{trace}\left[[E(S - S_{DL})]^T EE^T [E(S - S_{DL})] I_L \right] \qquad \because X - X_{DL} = ES - ES_{DL} = E(S - S_{DL}) \\
&= \mathrm{trace}\left[(S - S_{DL})^T E^T E E^T E (S - S_{DL}) I_L \right] \\
&= \mathrm{trace}\left[\underbrace{(S - S_{DL})^T}_{L\times N} \underbrace{I_N}_{N\times N} \underbrace{I_N}_{N\times N} \underbrace{(S - S_{DL})}_{N\times L} \underbrace{I_L}_{L\times L} \right] \qquad \because E^T E = I_N \\
&= \mathrm{trace}\left[(S - S_{DL})^T (S - S_{DL}) \right] \qquad \because \|A\|_F = \sqrt{\mathrm{trace}(A'A)} \\
&= \| S - S_{DL} \|_F^2 \quad \divideontimes
\end{aligned}$$