Final Review - ch2

For $C[a,b]$, example norms: $\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$,

$$\|f\|_1 = \int_a^b |f(x)|\,dx$$

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx\right)^{\frac{1}{2}}$$

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx\right)^{1/p}$$

For l_∞, l_∞-norm: $\forall a \in l_\infty$, $\|a\|_\infty = \sup_i |a_i|$

For l_p,

Definition of l_p-space: $\{a \mid \|a\|_p < +\infty\} \subset l_\infty$.

l_p-norm: $\forall$ Infinite sequence $a = (a_i)_{i\in\mathbb{Z}}$,

$$\|a\|_p = \left(\sum_{i=1}^{\infty} |a_i|^p\right)^{\frac{1}{p}}, \text{ where } p \geq 1.$$

Example: $a = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \\ \vdots \end{pmatrix}$, $\|a\|_\infty = 1$, $\|a\|_2 = \frac{\pi}{6}$, $\|a\|_1 = +\infty$, $\boxed{1 > f ③}$

$\therefore a \in l_\infty$, $a \in l_2$, $a \notin l_1$

Section 1.2. Case study: Clustering, k-means & k-medians

Clustering:

Let $C_i \in \{1, 2, \dots, k\}$ = x_i belongs to which group

i from 1 to N.

Let $G_j \subset \{1, 2, \dots, N\}$ = data points of group j

j from 1 to k

$\Rightarrow G_j = \{i \mid C_i = j\}$, $C_i = j \;\; \forall i \in G_j$

Interpretation:
If data point i belongs to group j

c係每個point屬於咩group

Gj = group j has what data point

Interpretation:
If data point i belongs to group j

c係每個point屬於咩group

G係每個group有咩point

G_j = group j has what data point
C_i = which group does this data point belongs to

Evaluation:
Within one specific group j, all vectors be close to representative vector z_j

Let

$$J_j = \sum_{i \in G_j} dist(x_i, z_j) = \sum_{i \in G_j} \|x_i - z_j\|_2^2$$

Altogether, we solve the following

姐係咁多group一齊都要最細

$$\min_{\substack{G_1, \dots, G_K \\ z_1, \dots, z_K}} J \iff \min_{\substack{G_1, \dots, G_K \\ z_1, \dots, z_K}} \sum_{j=1}^{K} J_j \iff \min_{\substack{G_1, \dots, G_K \\ z_1, \dots, z_K}} \sum_{j=1}^{K} \left(\sum_{i \in G_j} \|x_i - z_j\|_2^2 \right)$$

K-mean Algorithm:

Input: $x_1, x_2, \dots, x_N \in \mathbb{R}^n$

output: $c_1, c_2, \dots, c_N$ and z_j, $j = 1, \dots k$

Initialization: Initialize $z_1, z_2, \dots, z_K$ by choosing K vectors from $x_1, \dots x_N$ randomly

① step 1: Given $z_1, z_2, \dots, z_N$, compute

$$c_i = \arg\min_{j = \{1, 2, \dots, k\}} \|x_i - z_j\|_2 \quad , i = 1, 2, \dots N$$

define

$$G_j = \{ i \mid c_i = j \} \quad , j = 1, 2, \dots k$$

(人話: x_i is assigned to the group whose representative vector is the closet to x)

② step 2: Given $G_1, G_2, \dots, G_K$, compute

$$z_j = \frac{1}{|G_j|} \left(\sum_{i \in G_j} x_i \right)$$

(人話: compute z_j, the mean of all vectors in G_j)

Refer to pf 4 proof of clustering algorithm (K-mean)

K-median algorithm:

Replace
$$\min_{\substack{G_1,\cdots,G_K\\ z_1,\cdots,z_K}} \sum_{j=1}^{K}\left(\sum_{i\in G_j}\|x_i - z_j\|_2^2\right)$$
by
$$\min_{\substack{G_1,\cdots,G_K\\ z_1,\cdots,z_K}} \sum_{j=1}^{K}\left(\sum_{i\in G_j}\|x_i - z_j\|_1\right)$$

The numerical solver is

Step 1: Fix $z_1, \cdots, z_K$, solve
$$\min_{G_1,\cdots,G_K} \sum_{j=1}^{K}\left(\sum_{i\in G_j}\|x_i - z_j\|_1\right).$$
Similar to the discussion in k-means, the solution is
$$C_i = \operatorname*{argmin}_{j\in\{1,2,\cdots,K\}} \|x_i - z_j\|_1, \qquad i=1,2,\cdots,N.$$
and $G_j = \{ i \mid C_i = j\}$.

Step 2: Fix $G_1, G_2, \cdots, G_K$, solve
$$\min_{z_1,\cdots,z_K} \sum_{j=1}^{K}\left(\sum_{i\in G_j}\|x_i - z_j\|_1\right)$$

Similar to the discussion in k-means, it is decomposed into K sub problems
$$\min_{z_j} \sum_{i\in G_j}\|x_i - z_j\|_1 \qquad j=1,2,\cdots,K.$$
It is well known (Galileo) that the solution is
$$z_j = \operatorname*{median}_{i\in G_j}(x_i) \leftarrow \text{median of each entry.}$$
where $\operatorname*{median}_{i\in G_j}(x_i)$ takes component-wise median.

Median is more robust to outliers than k-mean
K-median is better than k-means if the dataset contains many outliers

---

Section 2.3 Limit and convergence in vector spaces

$\{x^{(k)}\}_{k\in\mathbb{N}} \in V$

Let $x \in V$. We say $\{x^{(k)}\}_{k\in\mathbb{N}}$ converges to x, denoted by $x^{(k)} \to x$, if
$$\lim_{k\to\infty}\|x^{(k)} - x\| = 0$$

驗證：要自己想兩支在V中的vector出來，sub進去norm然後計算norm,再limit k -> inf

Example: 1. $\mathbb{R}^n$ with $\|x\|_2$ is convergent to $\vec{0}$

2 $C[a,b]$ with $\|f\|_\infty$ is convergent to $g(x)$

Example: 1. $\mathbb{R}^n$ with $\|x\|_2$ is convergent

2 $C[0,1]$ with $\|f\|_\infty$ is convergent to $g(x)$

PS: 常用trick: $f^{(k)}(t) = \sin(2\pi kt)$, max of |sine function| = 1

special example: $a^{(k)} = \begin{pmatrix} \underbrace{1/k \\ \vdots \\ 1/k}_{k \text{ terms}} \\ 0 \\ \vdots \end{pmatrix} \in l_1, l_2, l_\infty$.

l_2, l_∞ convergent to 0. l_1 does not.

Recall: l(x) space mean that $\{a \text{ is an inf sequence} \mid \|a\|_x < +\infty\}$

(with a finite norm value)

We can check that $\|a^{(k)} - a\|_1 \to 1$, not 0

==> convergence / limits depends on norms

Special example 2:

: Consider $V=\{a \mid a \text{ is an infinite sequence}, \|a\|_1<+\infty\}$ with $\|\cdot\|_\infty$ norm

$\Leftarrow\Rightarrow$ $(V, \|\cdot\|_\infty)$

(反證技巧, 作 $a^{(k)}\in V$, $a\notin V$)

Let:

$$a^{(k)}=\begin{pmatrix}1\\ 1/2\\ \vdots\\ 1/k\\ 0\\ \vdots\end{pmatrix}\in V \qquad a=\begin{pmatrix}1\\ 1/2\\ 1/3\\ \vdots\\ 1/k\\ \vdots\\ \vdots\end{pmatrix}\notin V.$$

$$\|a^{(k)}\|_1=\sum_{i=1}^{k}\left(\tfrac{1}{i}\right)<+\infty \quad \text{as} \quad k<+\infty.$$

$$\|a\|_1=\sum_{i=1}^{k}\left(\tfrac{1}{i}\right)+\sum_{i=1}^{\infty}\left(\tfrac{1}{k}\right)=+\infty.$$

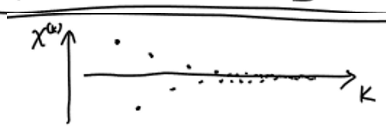
$$\lim_{k\to\infty}\|a^{(k)}-a\|_\infty=\lim_{k\to\infty}\left\|\begin{pmatrix}0\\ \vdots\\ 0\\ 1/(k+1)\\ 1/(k+2)\\ \vdots\end{pmatrix}\right\|_\infty=\lim_{k\to\infty}\frac{1}{k+1}=0$$

- This example shows that: The limit may not be in a normed vector space.
- If this happen, we call the normed vector space incomplete

Section 2.4: Completeness of normed vector spaces

Given a sequence in V with norm ||x||. Determine the sequence to be convergent

$$\boxed{\begin{array}{l}\text{Cauchy sequence: } \{x^{(k)}\}_{k\in\mathbb{N}} \text{ is called a cauchy sequence if}\\ \forall \varepsilon>0,\ \exists K \text{ s.t. } \forall k,l\geq K \quad \|x^{(k)}-x^{(l)}\|<\varepsilon\end{array}}$$



白話： 只要這個sequence 在某個閾值(K) 之後的k,l都保持$||x^{(k)}-x^{(l)}||$ 差距很小很小
就叫做cauchy sequence

Lemma: If $x^{(k)}\to x\in V$, then $\{x^{(k)}\}$ is a cauchy sequence.

(Pf5)

Note:

(Pf5)

Note: The reverse of the lemma is not true.

(Pf6)

A vector space V with norm $\|\cdot\|$ is complete if all cauchy sequence in V is convergent.

Banach space = a complete normed vector space

We can always complete a normed vector space by including all limit of Cauchy sequences

Example of complete normed vector spaces (Banach spaces)

- $\mathbb{R}^n$ with any norm.
- $\mathbb{R}^{n\times m}$ with any norm.
- Tensor space $\mathbb{R}^{m\times n\times l}$ with any norm
- $C[a,b]$ with $\|\cdot\|_\infty$ norm
- l_p with $p\geq 1$ and finite, l_∞

Example of incomplete normed vector spaces.

- Example 4 is the last section:

  $V=\left\{a \,\middle|\, \begin{matrix} a \text{ is a infinite sequence} \\ \|a\|_1 < +\infty \end{matrix}\right\}$ with norm $\|\cdot\|_\infty$ is incomplete

  The completion of this space is l_∞.
- $C[a,b]$ with p-norm, $p\neq +\infty$, $p\geq 1$, are incomplete.

Application: most of the time we will be dealing with iterative algorithms, we need to ensure the limit of the sequence converge to x \belongs to V
E.g. supervised learning want to find f \belongs c[a,b], so we don't want to use $||f^{(iter)} - f||_2$ 2 norm to measure distance, since $||f||_2$ in c[a,b] is not complete

2.4 Finite dimensional vector spaces
Properties:

1. ① For a <u>finite dimensional vector space</u> V, all norms are equivalent in the sense that:

每一個都可以被 constant bound住

For any two norms $\|\cdot\|_A$ and $\|\cdot\|_B$, there exists $C_1, C_2 > 0$ such that: $C_1\|a\|_A \le \|a\|_B \le C_2\|a\|_A \quad \forall a \in V$.

- Consequently, the limit of the same sequence under any norm is the same.

$$x^{(k)} \to x \text{ in } \|\cdot\|_A \iff x^{(k)} \to x \text{ in } \|\cdot\|_B$$

Note that the constants C_1, C_2 depend on V.

Example: Consider $\mathbb{R}^n$ and $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$

- $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent because

$$\|a\|_2 \le \|a\|_1 \le \sqrt{n}\,\|a\|_2$$

- $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are equivalent

$$\|a\|_\infty \le \|a\|_2 \le \sqrt{n}\,\|a\|_\infty$$

- $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are equivalent

$$\|a\|_\infty \le \|a\|_1 \le n\,\|a\|_\infty$$

Pf7: proof the above inequalities

② Any finite dimensional normed vector space is complete (i.e., they are Banach spaces)