Homework - TSNE

201800130112 赵子涵

1. 公式推导

$$\begin{cases} q_{ij} = q_{ji} = \dfrac{(1+\|y_i - y_j\|^2)^{-1}}{\sum_{k,l\neq k}(1+\|y_k - y_l\|^2)^{-1}} = \dfrac{E_{ij}^{-1}}{Z} \\ C = \sum\limits_{k,l\neq k} p_{lk} \log \dfrac{p_{lk}}{q_{lk}} \end{cases}$$

推导：

$$C = \sum_{k,l\neq k} p_{lk} \log \frac{p_{lk}}{q_{lk}}$$

$$= \sum_{k,l\neq k} (p_{lk} \log p_{lk} - p_{lk} \log q_{lk})$$

$$= \sum_{k,l\neq k} (p_{lk} \log p_{lk} - p_{lk} \log E_{kl}^{-1} + p_{lk} \log Z)$$

$$\frac{\partial C}{\partial y_i} = \sum_{k,l\neq k} \left(-p_{lk} \frac{\partial \log E_{kl}^{-1}}{\partial y_i} + p_{lk} \frac{\partial \log Z}{\partial y_i}\right)$$

对于 E_{kl}，显然有 $E_{kl} = E_{lk}$，$p_{lk} = p_{kl}$
当 k,l 中，$k=i$ 或 $l=i$，更换下标为 ij，
则有：

$$\sum_{k,l\neq k} -p_{lk} \frac{\partial \log E_{kl}^{-1}}{\partial y_i} = \sum_{j\neq i} -2\cdot p_{ij} \frac{\partial \log E_{ij}^{-1}}{\partial y_i}$$

$$\frac{\partial \log E_{ij}^{-1}}{\partial y_i} = \frac{1}{E_{ij}^{-1}} \times E_{ij}^{-2} \times (-2) \times (y_i - y_j) \times 1$$

$$\therefore \sum_{j\neq i} -2\cdot p_{ij} \frac{\partial \log E_{ij}^{-1}}{\partial y_i} = \sum_{j\neq i} 4 p_{ij} E_{ij}^{-1} (y_i - y_j)$$

$$\sum_{k,l\neq k} p_{lk} \frac{\partial \log Z}{\partial y_i} = \frac{\partial \log Z}{\partial y_i} \times \sum_{k,l\neq k} p_{lk}$$

$$= \frac{\partial \log Z}{\partial y_i}$$

$$= \frac{1}{Z} \times \sum_{k,l\neq k} \frac{\partial E_{k,l}^{-1}}{\partial y_i}$$

$$= 2 \times \frac{1}{Z} \times \sum_{j\neq i} E_{ij}^{-2} \times (-2) \times (y_i - y_j)$$

$$\because q_{ij} = q_{ji} = \frac{E_{ij}^{-1}}{Z}$$

$$\therefore Z = \frac{E_{ij}^{-1}}{q_{ij}}$$

$\therefore$ 上式 $= -4 \times \sum\limits_{j\neq i} q_{ij} \cdot E_{ij}^{-1} \cdot (y_i - y_j)$

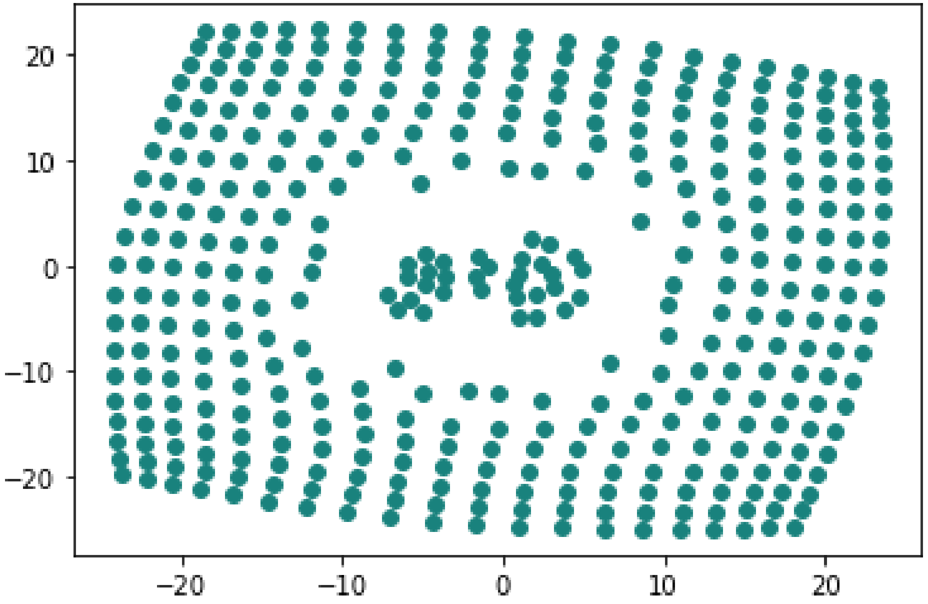
综上，

$$\frac{\partial C}{\partial y_i} = \sum_{j\neq i} 4(p_{ij} - q_{ij}) \times (1+\|y_i - y_j\|^2)^{-1} \times (y_i - y_j)$$

2. tsne参数与图像 —— 使用scikit-learn默认参数值

3D数据说明：

- 3D高斯
- 20 * 20，共400个数据
- 降维前已标准化

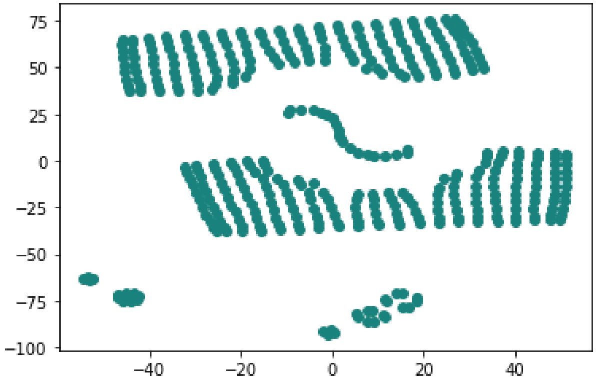


- n_components=2
- perplexity=30.0
- early_exaggeration=12.0
- learning_rate=200.0
- n_iter=1000
- n_iter_without_progress=300
- min_grad_norm=1e-07
- metric='euclidean'
- init='random'
- verbose=0
- random_state=0（默认值None，为保证每次结果相同设置为0）
- method='barnes_hut'
- angle=0.5
- n_jobs=None

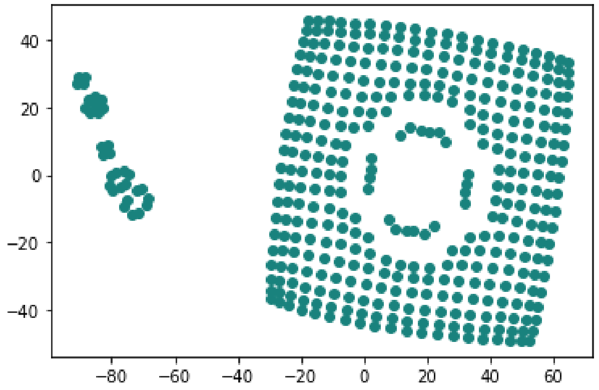
2. tsne参数与图像 —— 仅调整perplexity

- perplexity： number of nearest neighbors
- random_state=0（以保证每次生成的结果相同）

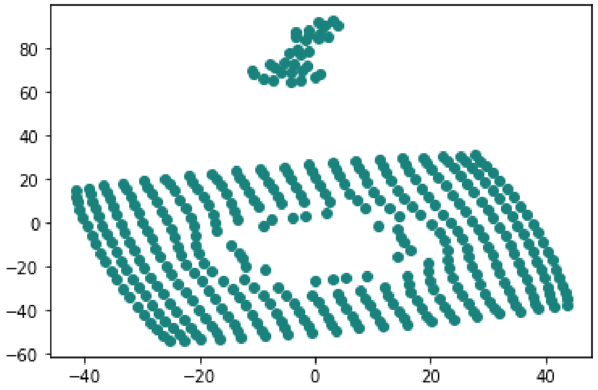
perplexity=5



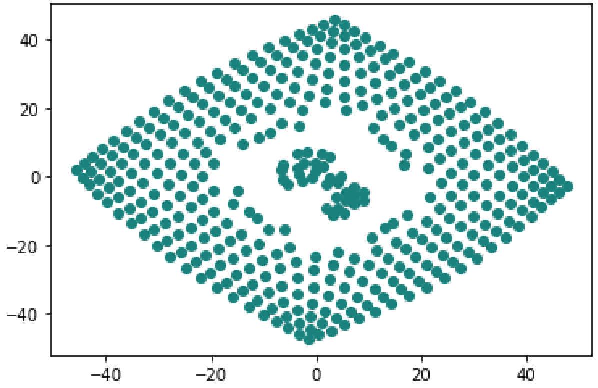
perplexity=10



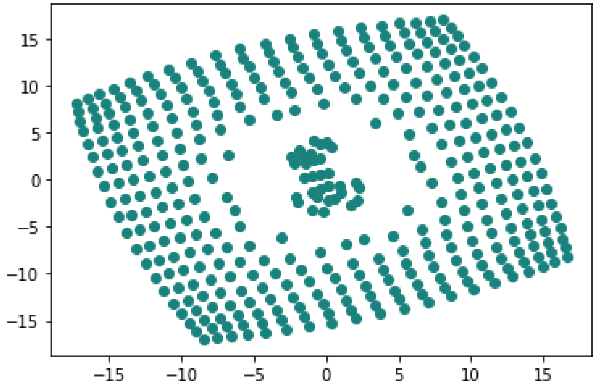
perplexity=15



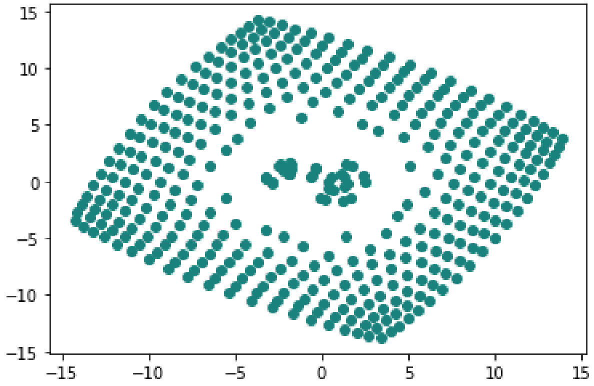
perplexity=20



perplexity=40



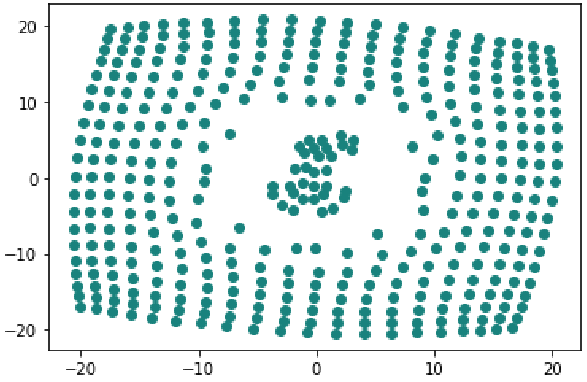
perplexity=50



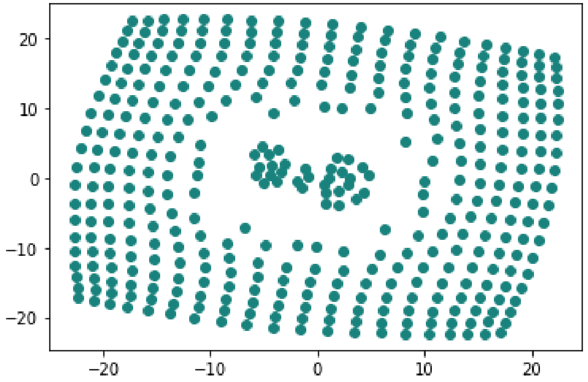
2. tsne参数与图像 —— 仅调整learning rate

- learning rate
- random_state=0（以保证每次生成的结果相同）

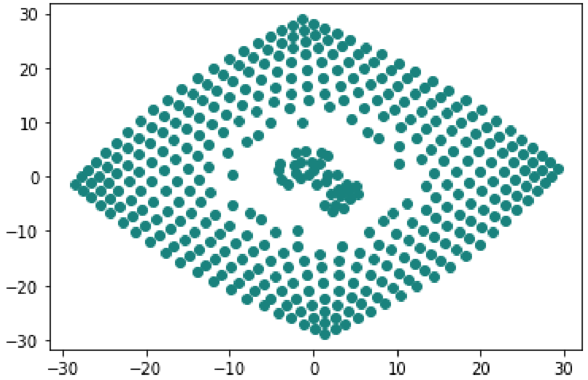
learning rate =10



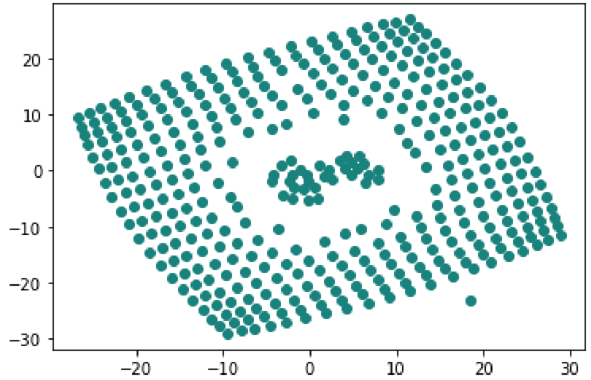
learning rate =50



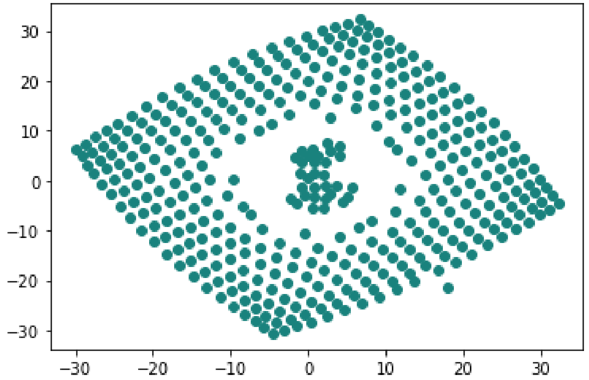
learning rate =100



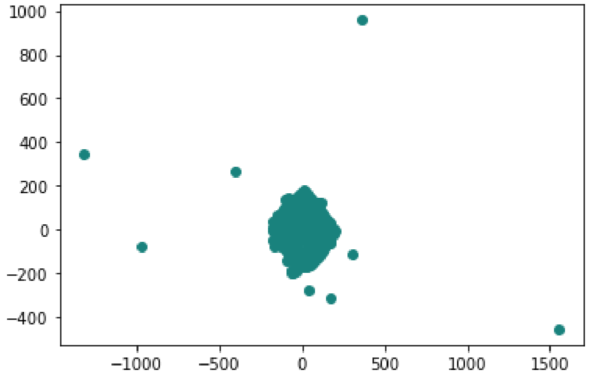
learning rate =500



learning rate =1000



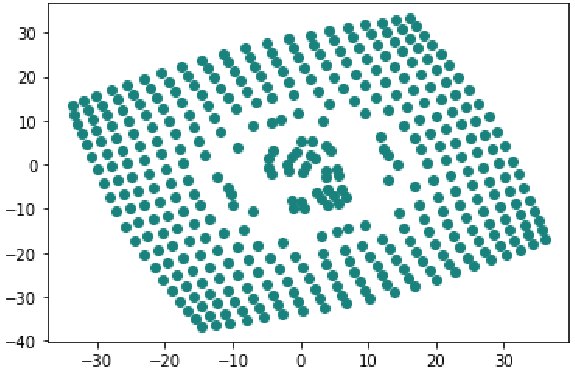
learning rate =1500



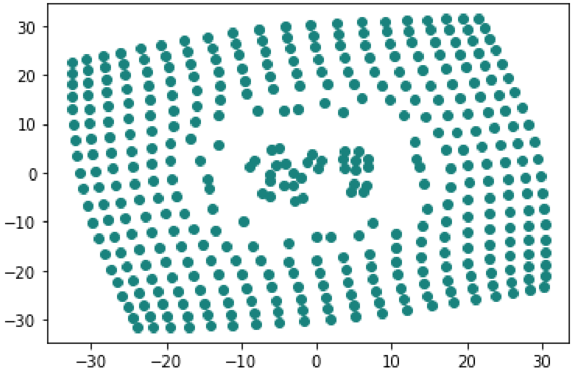
2. tsne参数与图像 —— 仅调整metric

- metric： calculating distance between instances in a feature array
- random_state=0（以保证每次生成的结果相同）

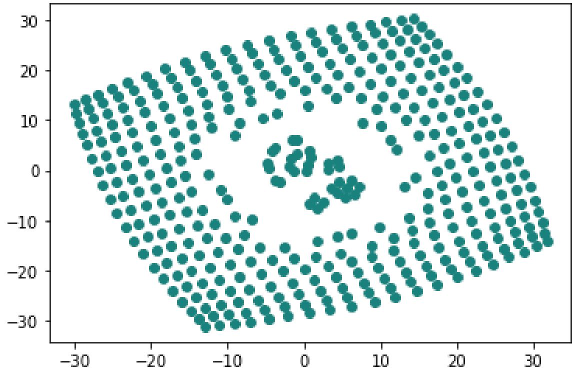
cityblock



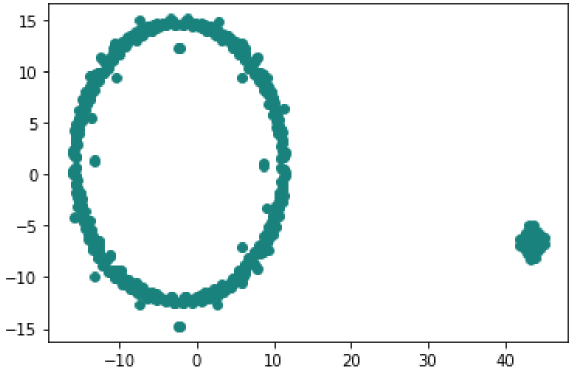
minkowski



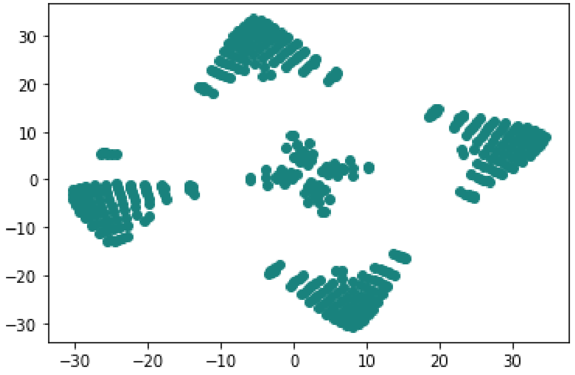
chebyshev



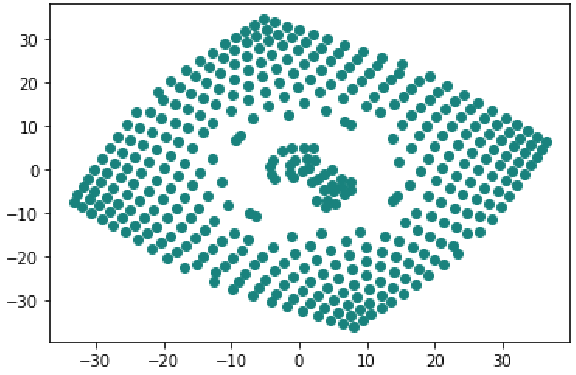
cosine



canberra



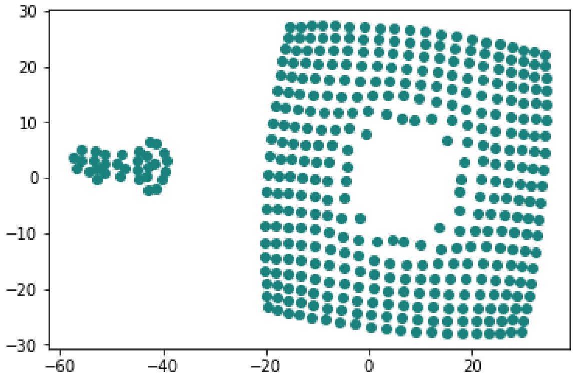
braycurtis



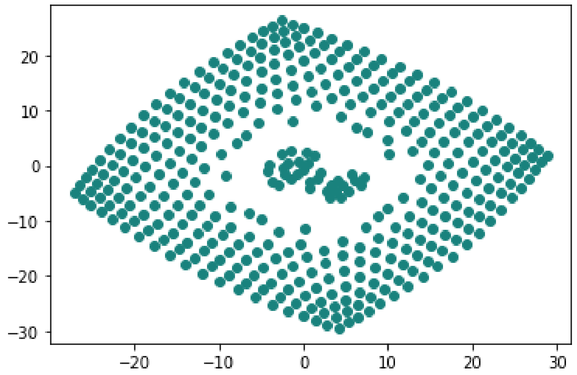
2. tsne参数与图像 —— 仅调整early_exaggeration

- early_exaggeration： Controls how tight natural clusters in the original space are in the embedded space and how much space will be between them.
- random_state=0（以保证每次生成的结果相同）

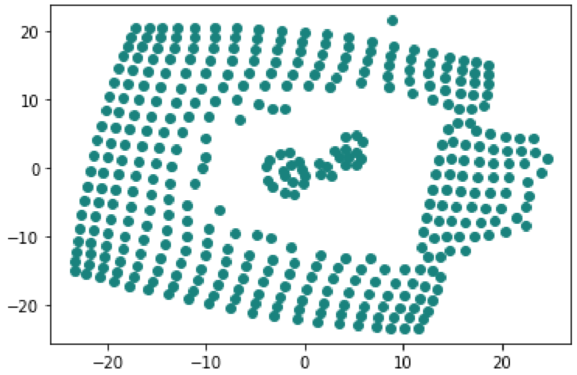
early_exaggeration=5



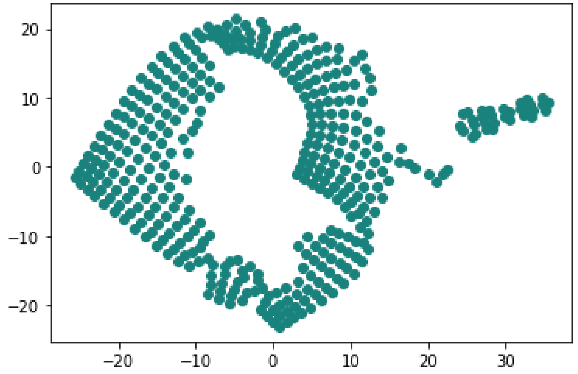
early_exaggeration=10



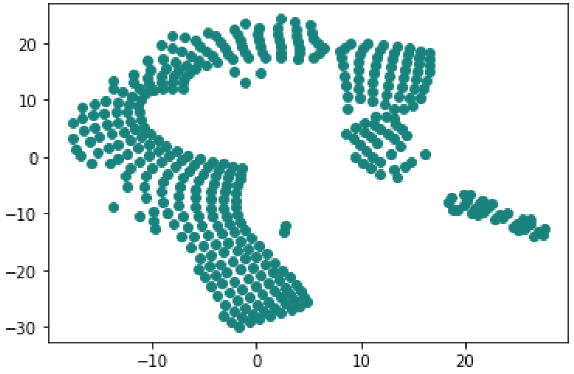
early_exaggeration=50



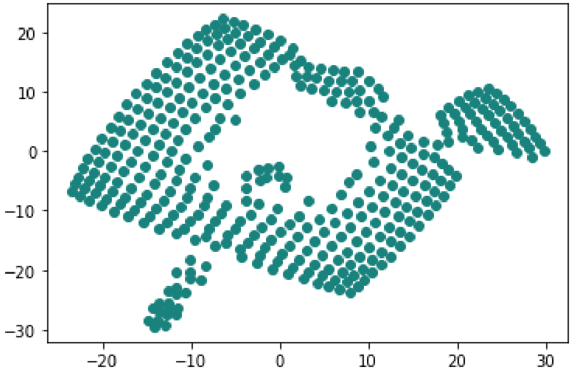
early_exaggeration=100



early_exaggeration=150



early_exaggeration=200



THANKS

201800130112 赵子涵