Diversity-aware sampling

If we have a kernel measuring similarity between any inputs, can define

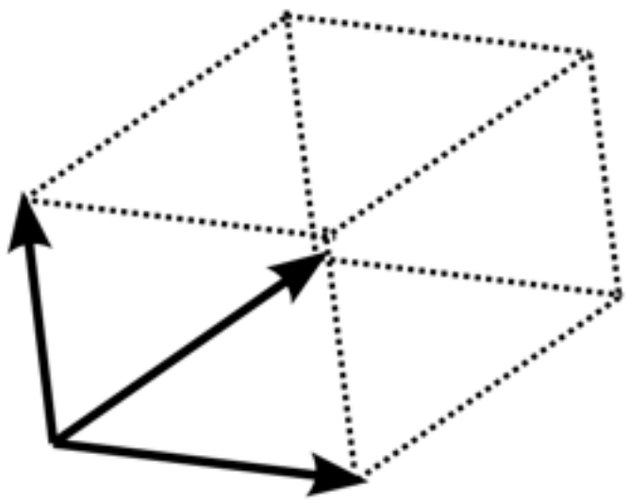
$$D(\{\theta_i\}_{i=1}^n) = \log\det\left(\begin{bmatrix} k(\theta_1,\theta_1) & \cdots & k(\theta_1,\theta_n) \\ \cdots & & \cdots \\ k(\theta_n,\theta_1) & \cdots & k(\theta_n,\theta_n) \end{bmatrix}\sigma^{-2} + \mathrm{I}\right)$$

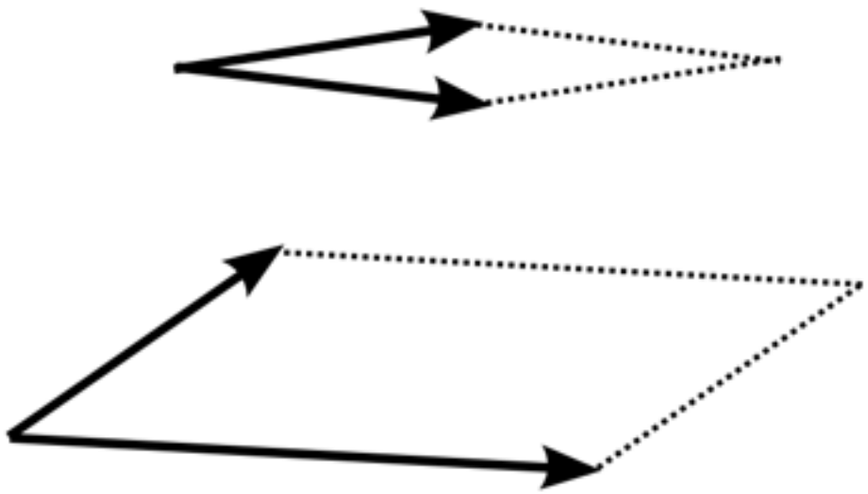
diversity metric

kernel

free parameter

identity matrix





[Kulesza&Taskar, 2013]

Diversity-aware sampling

If we have a kernel measuring similarity between any inputs, can define

$$D(\{\theta_i\}_{i=1}^n) = \log\det\left(\begin{bmatrix} k(\theta_1,\theta_1) & \cdots & k(\theta_1,\theta_n) \\ \cdots & & \cdots \\ k(\theta_n,\theta_1) & \cdots & k(\theta_n,\theta_n) \end{bmatrix}\sigma^{-2} + \mathrm{I}\right)$$

diversity metric

kernel

free parameter

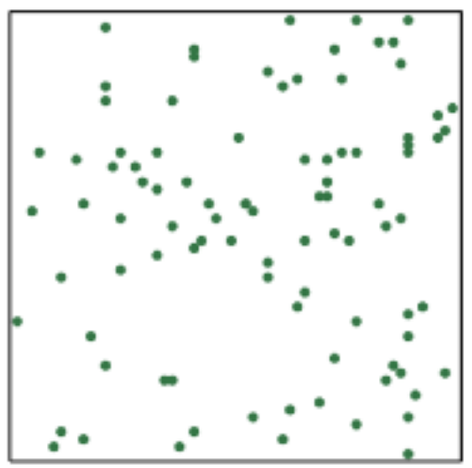
identity matrix

Generate an ordering of samples by greedily optimizing $D(\cdot)$

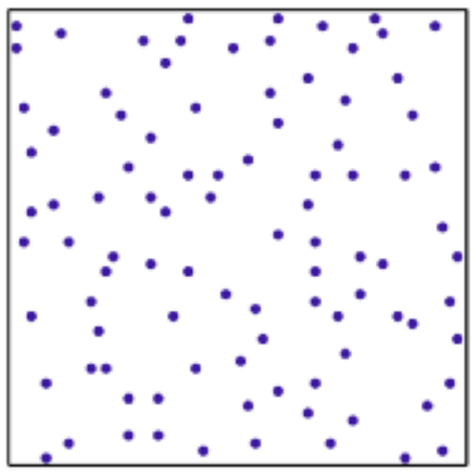
For $i = 1 \rightarrow n$

$$\theta_i = \text{argmax}_\theta\, D(\theta \cup \{\theta_j\}_{j=1}^{i-1})$$

[Kulesza&Taskar, 2013]



independent



diverse