CIL Series 3

The Nonstandard Deviations

March 17, 2015

Problem 1

1. $P(G) = P[\{BG, GB, GG\}] = 3/4$
2. $P(GG) = P[\{GG\}] = 1/4$
3. $P(GG|GX) = P[\{GG\}]/P[\{GB, GG\}] = 1/2$
4. $P(GG|G) = P[\{GG\}]/P[\{BG, GB, GG\}] = 1/3$
5. $P(GG|G) = P[\{GG\}]/P[\{BG, GB, GG\}] = 1/3$

Problem 2

- $p(\text{inf}) = (1/100)$
- $p(\text{pos}|\text{inf}) = (99/100)$
- $p(\text{neg}|\text{not inf}) = (99/100)$
- $p(\text{inf}|\text{pos}) = p(\text{pos}|\text{inf}) * p(\text{inf})/p(\text{pos})$
- $p(\text{pos}) = p(\text{pos}, \text{inf}) + p(\text{pos}, \text{not inf}) = p(\text{pos}|\text{inf}) * p(\text{inf}) + p(\text{pos}|\text{not inf}) * p(\text{not inf}) = (99/100) * (1/100) + (1/100) * (99/100) = (99/5000)$
- $p(\text{inf}|\text{pos}) = p(\text{pos}|\text{inf}) * p(\text{inf})/p(\text{pos}) = (99/100) * (1/100)/(99/5000) = 1/2$

Problem 3

1. - Optimal Assignment: Consider $\mathbf{z}_n$ for arbitrary n. By the hard assignment constraint we have $\sum_{k=1}^{K} z_{k,n} = 1$ with $z_{k,n} \in \{0, 1\}$. Line 2 dictates that $z_{k,n} = 1$ if $k = k^*$ where $k^* = \arg\min_{k \in \{1,\ldots,K\}} \|\mathbf{x}_n - \mathbf{u}_k\|_2^2$. Now consider the case where $z_{k',n} = 1$ and k' does not minimize $\|\mathbf{x}_n - \mathbf{u}_k\|_2^2$. Then $J_n = \sum_{k=1}^{K} z_{k,n} \|\mathbf{x}_n - \mathbf{u}_k\|_2^2 = \|\mathbf{x}_n - \mathbf{u}_{k'}\|_2^2 > \|\mathbf{x}_n - \mathbf{u}_{k^*}\|_2^2$. Hence this alternate assignment can not be optimal.

- Optimal Centroids: Lecture slide 12/38 already shows that

$$\nabla_{\mathbf{U}} J(\mathbf{U}, \mathbf{Z}) \stackrel{!}{=} 0 \implies \mathbf{u}_k^*(\mathbf{Z}) = \frac{\sum_{n=1}^N z_{k,n} \mathbf{x}_n}{\sum_{n=1}^N z_{k,n}}, \text{ if} \sum_{n=1}^N z_{k,n} > 0.$$

  Since $J(\mathbf{U}, \mathbf{Z})$ is convex when $\mathbf{Z}$ is fixed, this shows that the given centroids are optimal.

2. *Proof.*

$$\begin{aligned}
\|\mathbf{X} - \mathbf{U}\mathbf{Z}\|_F^2 &= \sum_{d=1}^D \sum_{n=1}^N (x_{dn} - \mathbf{U}\mathbf{Z}_{dn})^2 && \text{(definition of matrix multiplication)} \\
&= \sum_{d=1}^D \sum_{n=1}^N (x_{dn} - \sum_{k=1}^K u_{dk} z_{kn})^2 && \text{(definition of Frobenious norm)} \\
&= \sum_{d=1}^D \sum_{n=1}^N (x_{dn} - u_{dk^*})^2 && (z_{k^*n} = 1, z_{kn} = 0 \text{ for } k \neq k^*) \\
&= \sum_{d=1}^D \sum_{n=1}^N \sum_{k=1}^K z_{kn} (x_{dn} - u_{dk})^2 && (z_{k^*n} = 1, z_{kn} = 0 \text{ for } k \neq k^*) \\
&= \sum_{n=1}^N \sum_{k=1}^K \sum_{d=1}^D z_{kn} (x_{dn} - u_{dk})^2 && \text{(reordering of summations)} \\
&= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \sum_{d=1}^D (x_{dn} - u_{dk})^2 && \text{(distributivity)} \\
&= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \|\mathbf{x}_n - \mathbf{u}_k\|^2 && \text{(definition of vector 2-norm)}
\end{aligned}$$

□