Computer Vision I: Low-Middle Level Vision Homework Exercise #2

(total 15 points)

Due: November 28th 11:59 PM.

2100017727 Keran Wang

Question 1. (Minimax entropy learning, 3 points).

Answer:

1.
$$Z = \int_{\mathbf{I}} \exp\left\{-\sum_{i}^{K} < \lambda_i, H_i(\mathbf{I}) >\right\}\mathbf{dI}$$

Thus,
$$\frac{\partial \log Z}{\partial \lambda_i} = \frac{1}{Z}\frac{\partial Z}{\partial \lambda_i} = -E_p[H_i(\mathbf{I})]$$

2. By definition, $\ell(\Theta) = -\log p(\mathbf{I}, \Theta)$, which means
$$\ell(\Theta) = -\log p(\mathbf{I}, \Theta) = -\log Z - \sum_{i=1}^{K} < \lambda_i, H_i(\mathbf{I}) >$$

Due to the nature of inner product, the second term vanishes in the second derivative.
$$\begin{aligned}
\frac{\partial^2 \ell(\Theta)}{\partial \lambda_i \partial \lambda_j} &= -\frac{\partial^2 \log Z}{\partial \lambda_i \partial \lambda_j} \\
&= \frac{1}{Z^2}\frac{\partial Z}{\partial \lambda_i}\frac{\partial Z}{\partial \lambda_j} - \frac{1}{Z}\frac{\partial^2 Z}{\partial \lambda_i \partial \lambda_j} \\
&= E_p[H_i(\mathbf{I})]E_p[H_j(\mathbf{I})] - E_p[H_i(\mathbf{I})H_j(\mathbf{I})] \\
&= -E_p[(H_i(\mathbf{I}) - h_i)(H_j(\mathbf{I}) - h_j)]
\end{aligned}$$

3. The left hand side of this equation is
$$\begin{aligned}
LHS &= E_f[\log f] - E_f[\log p] - E_f[\log f] + E_f[\log p_+] \\
&= E_f[\log p_+] - E_f[\log p] \\
&= E_f[-\log Z - \sum_{\alpha=1}^{K} < \lambda_\alpha^*, H_\alpha(\mathbf{I}) > - < \lambda_+, H_+(\mathbf{I}) >] - E_f[-\log Z - \sum_{\alpha=1}^{K} < \lambda_\alpha, H_\alpha(\mathbf{I}) >] \\
&= -\log Z - \sum_{\alpha=1}^{K} < \lambda_\alpha^*, E_f[H_\alpha(\mathbf{I})] > - < \lambda_+, E_f[H_+(\mathbf{I})] > + \log Z - \sum_{\alpha=1}^{K} < \lambda_\alpha, E_f[H_\alpha(\mathbf{I})] > \\
&= -\sum_{\alpha=1}^{K} < \lambda_\alpha^*, E_{p_+}[H_\alpha(\mathbf{I})] > - < \lambda_+, E_{p_+}[H_+(\mathbf{I})] > - \sum_{\alpha=1}^{K} < \lambda_\alpha, E_{p_+}[H_\alpha(\mathbf{I})] > \\
&= E_{p_+}[\log p_+] - E_{p_+}[\log p] = KL(p_+ || p)
\end{aligned}$$

Question 2. (Learning by information projection, 3 points).

Answer:

1. Here, we are searching for the solution for the constrained problem:

$$p = \arg\min_{p \in \Omega_i} \int p(\mathbf{I}) \log \frac{p(\mathbf{I})}{q(\mathbf{I})} \mathrm{d}\mathbf{I}$$

with constraint

$$\int p(\mathbf{I}) H_i(I) \mathrm{d}\mathbf{I} = h_i$$

and

$$\int p(\mathbf{I}) \mathrm{d}\mathbf{I} = 1$$

Using the Lagrarian method, we are optimizing the function:

$$\mathcal{L} = p(\mathbf{I}) \log \frac{p(\mathbf{I})}{q(\mathbf{I})} + \sum_{i=1}^{K} \lambda_i p(\mathbf{I}) H_i(I) + \lambda_0 p(\mathbf{I})$$

We have elimiated the constants useless for optimization. Thus, we have

$$\frac{\partial \mathcal{L}}{\partial p} = \log \frac{p(\mathbf{I})}{q(\mathbf{I})} + 1 + \sum_{i=1}^{K} \lambda_i H_i(I) + \lambda_0 = 0$$

Correspondingly,

$$p(\mathbf{I}) = \exp(-1 - \lambda_0 - \sum_{i=1}^{K} \lambda_i H_i(\mathbf{I})) q(\mathbf{I})$$

2. The left hand side of this equation is

$$\begin{aligned} LHS &= E_f[\log f] - E_f[\log q] - E_f[\log f] + E_f[\log p] \\ &= E_f[\log p] - E_f[\log q] \end{aligned}$$

Due to the equation of the expect for f and q, we have

$$\begin{aligned} LHS &= E_p[\log p] - E_p[\log q] \\ &= KL(p||q) \end{aligned}$$

3. When $q(\mathbf{I})$ is a uniform distribution, the expression can be rewritten as

$$\begin{aligned} p &= \arg\min_{p \in \Omega_i} \int p(\mathbf{I}) \log \frac{p(\mathbf{I})}{q(\mathbf{I})} \mathrm{d}\mathbf{I} \\ &= \arg\min_{p \in \Omega_i} \int p(\mathbf{I}) \log p(\mathbf{I}) \mathrm{d}\mathbf{I} \\ &= -\arg\max_{p \in \Omega_i} \int p(\mathbf{I}) \log p(\mathbf{I}) \mathrm{d}\mathbf{I} \end{aligned}$$

Question 3. (Information projection, 3 points)

Answer:

1. Similar to Quesion 1, here, β_K is λ_i, and q_{K-1} is p, thus,

$$\frac{\partial \log z_K}{\partial \beta_K} = E_{q_{K-1}}[h_K(x)]$$

2. Similarly, we have

$$\frac{\partial^2 \log Z_K}{\partial \beta_i \partial \beta_j} = E_{q_0}[(h_i(x) - h_i^{obs})(h_j(x) - h_j^{obs})]$$

3. Leveraging the equation between $E_f[\cdot]$ and $E_{q_{K+1}}[\cdot]$, the proof is similar.

Question 4. (Typical set, 3 points)

Answer:

1. Given N, the coin has qN times head up, the number is

$$\#\Omega(q) = \binom{N}{qN} = \frac{N!}{qN!(N-qN)!}$$

We regard N as a really big number, we have stirling's approximation

$$\#\Omega(q) = \frac{1}{\sqrt{2\pi Nq(1-q)} q^{qN}(1-q)^{(1-q)N}}$$

Substitute q with 0.2 and 0.5

$$\#\Omega(0.2) = \frac{1}{\sqrt{0.32\pi N} 0.2^{0.2N} 0.8^{0.8N}}$$

$$\#\Omega(0.5) = \frac{1}{\sqrt{\frac{\pi N}{2}}} 2^N$$

2.

$$p(S_N) = p^{qN}(1-p)^{(1-q)N}$$

$$p(\Omega(q)) = p(S_N)\#\Omega(q) = \frac{p^{qN}(1-p)^{(1-q)N}}{\sqrt{2\pi Nq(1-q)} q^{qN}(1-q)^{(1-q)N}}$$

3. When $p \neq q$, one of the two fractions $\frac{q}{p}$ and $\frac{p}{q}$ is lower than 1, and with the effect of infinite exponent, the probablity of being obeserved will converge to zero swiftly. So, we can observe it if and only if $p = q$.

Question 5. (Poisson Distribution and Random Sampling, 3 points)

Answer:

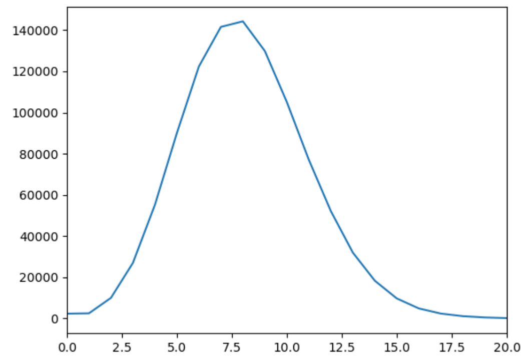

Figure 1: $\lambda = 0.1,\ L = 10$

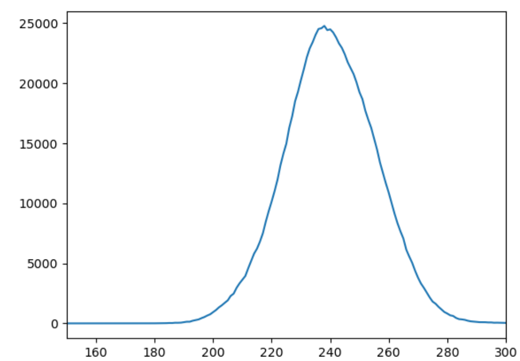

Figure 2: $\lambda = 0.1,\ L = 50$

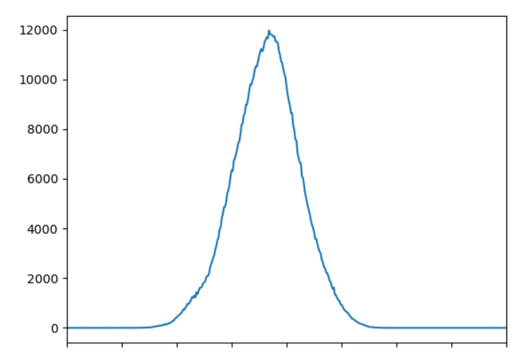

Figure 3: $\lambda = 0.1,\ L = 100$

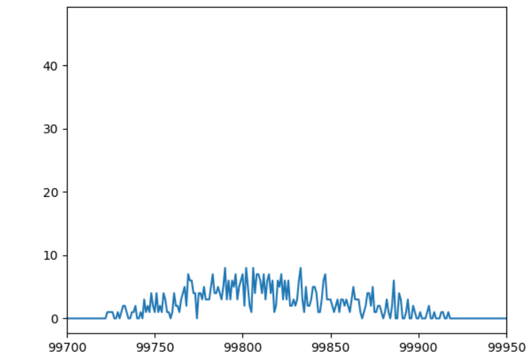

Figure 4: $\lambda = 0.1$, $L = 1000$

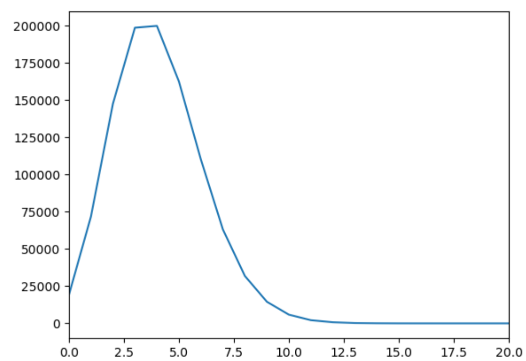

Figure 5: $\lambda = 0.05$, $L = 10$

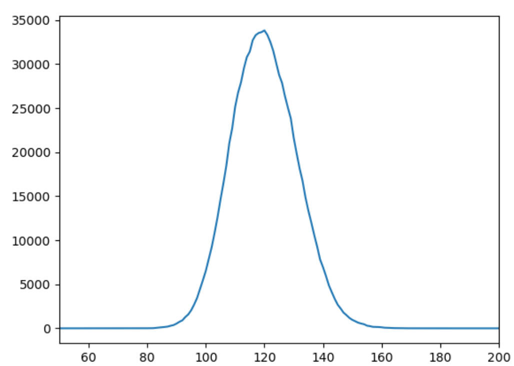

Figure 6: $\lambda = 0.05$, $L = 50$

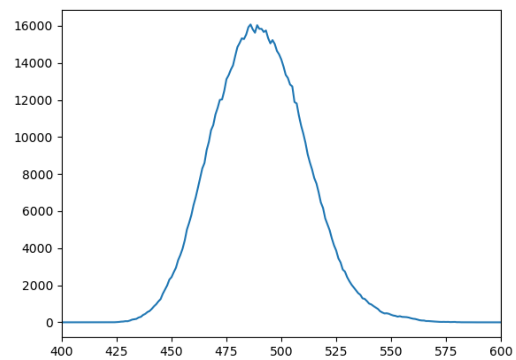

Figure 7: $\lambda = 0.05,\ L = 100$

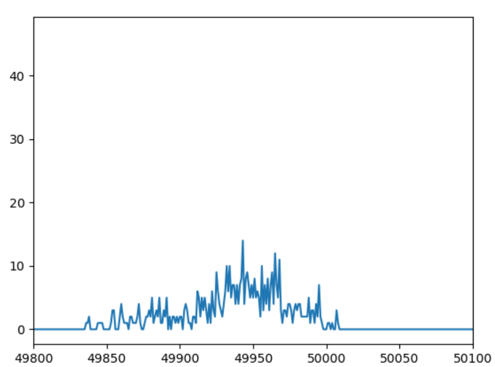

Figure 8: $\lambda = 0.05,\ L = 1000$

1.

2. The probability for one point in a $l \times l$ square is $\frac{l^2}{n^2}$, and the total probability is binom distribution:

$$P_s(k) = \binom{N}{k} (\frac{l^2}{n^2})^k (1 - \frac{l^2}{n^2})^{N-k}$$

When $\frac{l^2}{n^2}$ is very small, we can have an approximation:

$$(1 - \frac{l^2}{n^2})^{N-k} = \exp{(-\frac{(N-k)l^2}{n^2})}$$

So, we got

$$P_s(k) = \frac{(N\frac{l^2}{n^2})^k}{k!} e^{-\frac{Nl^2}{n^2}}$$

Notice that $N = \lambda l^2$, this is exact the Poisson Distribution