ESE 650, Spring 2020

Problem Set 0

Lyuheng Yuan[lyuheng@seas],
Collaborators: Lyuheng Yuan [lyuheng@seas]

January 28, 2020

Solution 1 (Time spent: 2 hours)**.** X and Y are N-dimensional Guassian distribution variables, every dimension of X or Y is a 1-dimensional Guassian distribution. First we prove the sum of two 1-dimensional Guassian distribution variables is also a Guassian, i.e.

$$X^{(1)} + Y^{(1)} \sim N(\mu_1^{(1)} + \mu_2^{(1)}, {\sigma_1^{(1)}}^2 + {\sigma_2^{(1)}}^2)$$

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(z-x-\mu_2)^2}{2\sigma_2^2}\right] \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right] dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sigma_1\sigma_2} \exp\left[-\frac{\sigma_1^2(z-x-\mu_2)^2 + \sigma_2^2(x-\mu_1)^2}{2\sigma_1^2\sigma_2^2}\right] dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sigma_1\sigma_2} \exp\left[-\frac{\sigma_1^2(z^2+x^2+\mu_2^2-2xz-2z\mu_2+2x\mu_2) + \sigma_2^2(x^2+\mu_1^2-2x\mu_1)}{2\sigma_2^2\sigma_1^2}\right] dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sigma_1\sigma_2} \exp\left[-\frac{x^2(\sigma_1^2+\sigma_2^2) - 2x(\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1) + \sigma_1^2(z^2+\mu_2^2-2z\mu_2)+\sigma_2^2\mu_1^2}{2\sigma_2^2\sigma_1^2}\right] dx$$

Assume $\sigma_3 = \sqrt{{\sigma_1}^2 + {\sigma_2}^2}$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_3} \frac{1}{\sqrt{2\pi}\frac{\sigma_1\sigma_2}{\sigma_3}} \exp\left[-\frac{x^2 - 2x\frac{\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1}{\sigma_3^2} + \frac{\sigma_1^2(z^2+\mu_2^2-2z\mu_2)+\sigma_2^2\mu_1^2}{\sigma_3^2}}{2\left(\frac{\sigma_1\sigma_2}{\sigma_3}\right)^2}\right] dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_3} \frac{1}{\sqrt{2\pi}\frac{\sigma_1\sigma_2}{\sigma_3}} \exp\left[-\frac{\left(x - \frac{\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1}{\sigma_3^2}\right)^2 - \left(\frac{\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1}{\sigma_3^2}\right)^2 + \frac{\sigma_1^2(z-\mu_2)^2+\sigma_2^2\mu_1^2}{\sigma_3^2}}{2\left(\frac{\sigma_1\sigma_2}{\sigma_3}\right)^2}\right] dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_3} \exp\left[-\frac{\sigma_3^2\left(\sigma_1^2(z-\mu_2)^2+\sigma_2^2\mu_1^2\right)-\left(\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1\right)^2}{2\sigma_3^2\left(\sigma_1\sigma_2\right)^2}\right] \frac{1}{\sqrt{2\pi}\frac{\sigma_1\sigma_2}{\sigma_3}} \exp\left[-\frac{\left(x-\frac{\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_2}{\sigma_3^2}\right)^2}{2\left(\frac{\sigma_1\sigma_2}{\sigma_3}\right)^2}\right.$$

$$= \frac{1}{\sqrt{2\pi}\sigma_3} \exp\left[-\frac{(z-(\mu_1+\mu_2))^2}{2\sigma_3^2}\right] \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\frac{\sigma_1\sigma_2}{\sigma_3}} \exp\left[-\frac{\left(x-\frac{\sigma_1^2(z-\mu_2)+\sigma_2^2\mu_1}{\sigma_3^2}\right)^2}{2\left(\frac{\sigma_1\sigma_2}{\sigma_3}\right)^2}\right] dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma_3} \exp\left[-\frac{(z-(\mu_1+\mu_2))^2}{2\sigma_3^2}\right]$$

Now we proved that every dimension of $Z = X + Y$, say $z_1, z_2, \cdots, z_n$ is Guassian distributions. Now we want to prove $[z_1, z_2, \cdots, z_n]^T$ is also a N-dimensional Guassian distribution. It's apparently correct because the variance of X and Y are diagonal matrix which means every dimension is independent from other dimensions. The mean and variance in every dimension is μ_n and σ_n respectively.

$$f(x) = \frac{1}{(\sqrt{2\pi})^d\sigma_1\sigma_2\cdots\sigma_n}\mathbf{exp}\{-\frac{1}{2}[(\frac{x_1-\mu_1}{\sigma_1})^2+(\frac{x_2-\mu_2}{\sigma_2})^2+\cdots+(\frac{x_d-\mu_1}{\sigma_d})^2]\}$$

$$= \frac{1}{(\sqrt{2\pi})\sigma_1}\mathbf{exp}\{-\frac{1}{2}(\frac{x_1-\mu_1}{\sigma_1})^2\}\frac{1}{(\sqrt{2\pi})\sigma_2}\mathbf{exp}\{-\frac{1}{2}(\frac{x_2-\mu_2}{\sigma_2})^2\}\cdots\frac{1}{(\sqrt{2\pi})\sigma_n}\mathbf{exp}\{-\frac{1}{2}(\frac{x_n-\mu_n}{\sigma_n})^2\}$$

$$d^2(x,\mu) = (\frac{x_1-\mu_1}{\sigma_1})^2+(\frac{x_2-\mu_2}{\sigma_2})^2+\cdots+(\frac{x_n-\mu_n}{\sigma_n})^2$$

$$= [x_1-\mu_1, x_2-\mu_2, \cdots, x_n-\mu_n]\begin{bmatrix}\frac{1}{\sigma_1^2}\ 0\cdots\ 0\\ 0\ \frac{1}{\sigma_2^2}\cdots\ 0\\ \vdots\ \vdots\ \ddots\ \vdots\\ 0\ \cdots\ \ \frac{1}{\sigma_n^2}\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\\ \vdots\\ x_n-\mu_n\end{bmatrix} = (X-\mu)^T\Sigma^{-1}(X-\mu)$$

$$f(x) = \frac{1}{(\sqrt{2\pi})^n(\sigma_1\sigma_2\cdots\sigma_d)}\mathbf{exp}\{-\frac{1}{2}d(x,\mu)^2\}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}\left|\Sigma\right|^{\frac{1}{2}}}\mathbf{exp}\{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\}$$

Now we have proved $[z_1, z_2, \cdots, z_n]^T$ is also a N-dimensional Guassian distribution. We can get the mean and variance according to this theorem.

$$\mathbf{E}[X+Y] = \mathbf{E}[X]+\mathbf{E}[Y] = \mu_1+\mu_2$$

$$\mathbf{Var}[X+Y] = \mathbf{Var}[X]+\mathbf{Var}[Y] = ({\sigma_1}^2+{\sigma_2}^2)I$$

Hence, $X+Y \sim \mathbf{N}(\mu_1+\mu_2, ({\sigma_1}^2+{\sigma_2}^2)I)$

Solution 2 (Time spent: 30 minutes)**.** Let us assume the prior probability is

$$\mathbf{P}(X=0)=\mathbf{P}(X=1)=0.5$$

The probability of the door is open when sensor detects that the door is open is

$$\mathbf{P}(X=0|Y=0)$$

According to Bayes Rule,

$$\mathbf{P}(X=0|Y=0)=\frac{\mathbf{P}(Y=0|X=0)\mathbf{P}(X=0)}{\mathbf{P}(Y=0|X=0)\mathbf{P}(X=0)+\mathbf{P}(Y=0|X=1)\mathbf{P}(X=1)}$$

$$=\frac{0.8\times 0.5}{0.8\times 0.5+0.2\times 0.5}=0.8$$

If we take multiple measurements, The probability should be

$$\mathbf{P}(X=0|Y_1=0,Y_2=0)$$

According to the Bayes Rule,

$$\mathbf{P}(X=0|Y_1=0,Y_2=0)=\frac{\mathbf{P}(Y_2=0|X=0)\mathbf{P}(X=0|Y_1=0)}{\mathbf{P}(Y_2=0|X=0)\mathbf{P}(X=0|Y_1=0)+\mathbf{P}(Y_2=0|X=1)\mathbf{P}(X=1|Y_1=0)}$$

$$=\frac{0.8\times 0.8}{0.8\times 0.8+0.2\times 0.2}=\frac{16}{17}=0.9412$$

We can see that the posteriori probability increases if we take multiple measurements.

Solution 3 (Time spent: 1.5 hours)**.** According to the linear system equation

$$x(1) = Ax(0) + Bu(0) + \xi(t)$$

Then

$$\mathbf{E}[x(1)] = \mathbf{E}[Ax(0) + Bu(0) + \xi(t)]$$

$$= \mathbf{E}[Ax(0)] + \mathbf{E}[Bu(0)] + \mathbf{E}[\xi(t)]$$

Since, $\xi \sim \mathbf{N}(0, \Sigma)$ and $x(0) \sim \mathbf{N}(0, I)$

$$\mathbf{E}[Ax(0)] + \mathbf{E}[Bu(0)] + \mathbf{E}[\xi(t)] = A\mathbf{E}[x(0)] + \mathbf{E}[Bu(0)] + \mathbf{E}[\xi(t)]$$

$$= 0 + Bu(0) + 0 = Bu(0)$$

For variance,

$$\mathbf{Var}[x(1)] = \mathbf{Var}[Ax(0) + Bu(0) + \xi(t)]$$

$$= \mathbf{Var}[Ax(0)] + \mathbf{Var}[\xi(t)]$$

$$= A\mathbf{Var}[x(0)]A^T + \mathbf{Var}[\xi(t)] = AA^T + \Sigma$$

So, $x(1) \sim \mathbf{N}(Bu(0), AA^T + \Sigma)$
$u(t)$ stabilize the system because $u(t)$ prevents the system from outputting a infinite number since $u(t)$ decides the mean of the system.

When t is close to positive infinity, $x(t+1) = x(t)$. Now we put this equation into the linear system

$$(I - A)x(t) = Bu(t) + \xi(t)$$

$$x(t) = (I - A)^{-1}(Bu(t) + \xi(t))$$

$$\lim_{t \to \infty} Var[x(t)] = (I - A)^{-1}\Sigma$$

The variance of $x(t)$ or $x(1)$ doesn't depend on $u(t)$ because $u(t)$ is just a control input which effects the mean of the system. It's a constant vector which doesn't have variance.

Solution 4 (Time spent: 1.5 hours)**.**

Solution 5 (Time spent: 3 hours)**.**

(i) Suppose t_i refers to the total number of bets if Jarvis have i coins at the beginning. P_{t_i} refers to $\mathbf{P}(X_{t_i} = n)$ which is winning probability, So we have

$$P_i = pP_{i+1} + qP_{i-1}$$

Change P_i into $pP_i + qP_i$, we get

$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1})$$

$$= (\frac{q}{p})^i (P_1 - P_0) = (\frac{q}{p})^i P_1$$

So

$$P_{i+1} = P_1 \sum_{k=0}^{i} (\frac{q}{p})^k = P_1 \frac{1 - (\frac{q}{p})^{i+1}}{1 - \frac{q}{p}}$$

Put $P_n = 1$ into the equation above.

$$P_1 \frac{1 - (\frac{q}{p})^n}{1 - \frac{q}{p}} = 1$$

We get

$$P_1 = \frac{1 - \frac{q}{p}}{1 - (\frac{q}{p})^n}$$

Hence, P_i is

$$P_i = \frac{1 - (\frac{q}{p})^i}{1 - (\frac{q}{p})^n}$$

According to $X_0 = m$, the probability of losing the game is

$$\mathbf{P}(win) = 1 - \frac{1 - (\frac{q}{p})^m}{1 - (\frac{q}{p})^n}$$

(ii) Suppose E_i refers to the total number of bets if Jarvis have i coins at the beginning. So we have

$$E_i = 1 + pE_{i+1} + qE_{i-1}$$

Change E_i into $pE_i + qE_i$, we get

$$E_{i+1} - E_i + \frac{1}{p-q} = \frac{q}{p}(E_i - E_{i-1} + \frac{1}{p-q})$$

$$= (\frac{q}{p})^i (E_1 - E_0 + \frac{1}{p-q}) = (\frac{q}{p})^i (E_1 + \frac{1}{p-q})$$

$$E_{i+1} + (i+1)\frac{1}{p-q} = \sum_{k=0}^{i} (\frac{q}{p})^k (E_1 + \frac{1}{p-q})$$

$$= (E_1 + \frac{1}{p-q}) \frac{1 - (\frac{q}{p})^{i+1}}{1 - \frac{q}{p}}$$

Put $E_n = 0$ into the equation above.

$$E_n = (E_1 + \frac{1}{p-q}) \frac{1 - (\frac{q}{p})^n}{1 - \frac{q}{p}} - \frac{n}{p-q}$$

$$E_1 = \frac{n}{p-q} \frac{1 - \frac{q}{p}}{1 - (\frac{q}{p})^n} - \frac{1}{p-q}$$

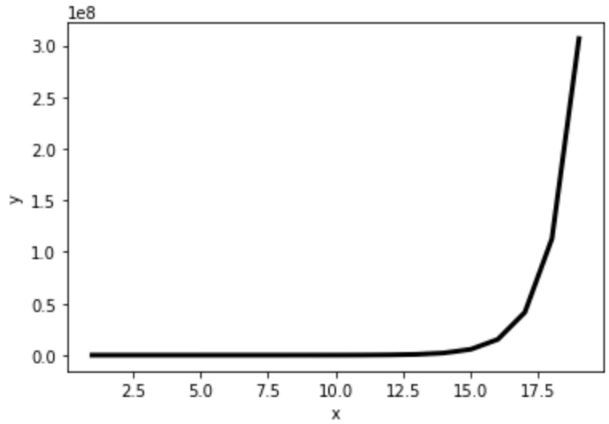
So

$$E_i = \frac{n}{p-q} \frac{1 - (\frac{q}{p})^i}{1 - (\frac{q}{p})^n} - \frac{i}{p-q}$$

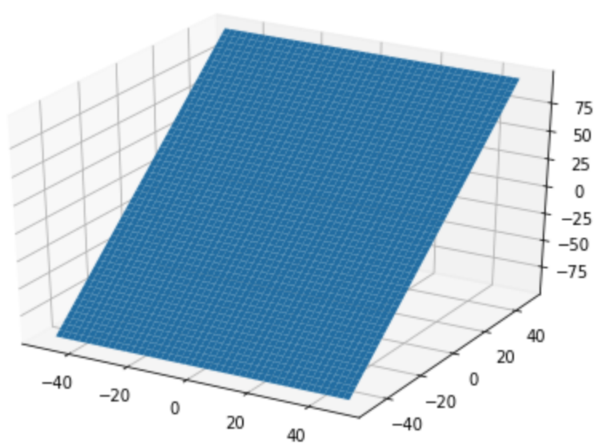
Now we get the expectation of E_m

$$E_m = \frac{n}{p-q} \frac{1 - (\frac{q}{p})^m}{1 - (\frac{q}{p})^n} - \frac{m}{p-q}$$

Solution 6 (Time spent: 3 hours)**.** For 1D function $y = e^x$, partial gradient is following



For 2D function $y = x^2 + y^2$, partial gradient w.r.t x is following



For 2D function $y = x^2 + y^2$, partial gradient w.r.t y is following