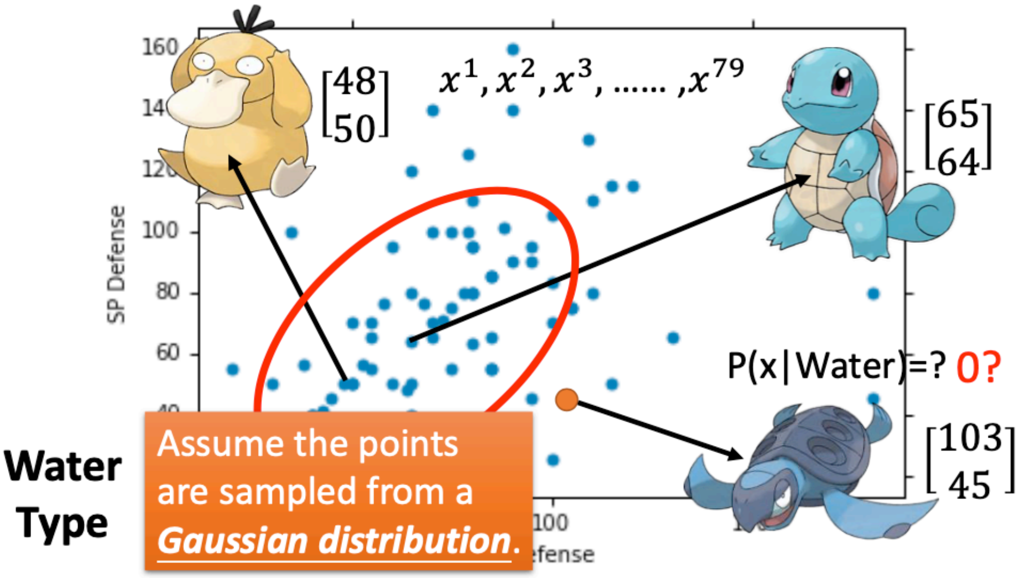
考虑到宝可梦的两个属性（**Defense**、**SP Defense**），将输入的宝可梦进行属性分类（Water、Normal）



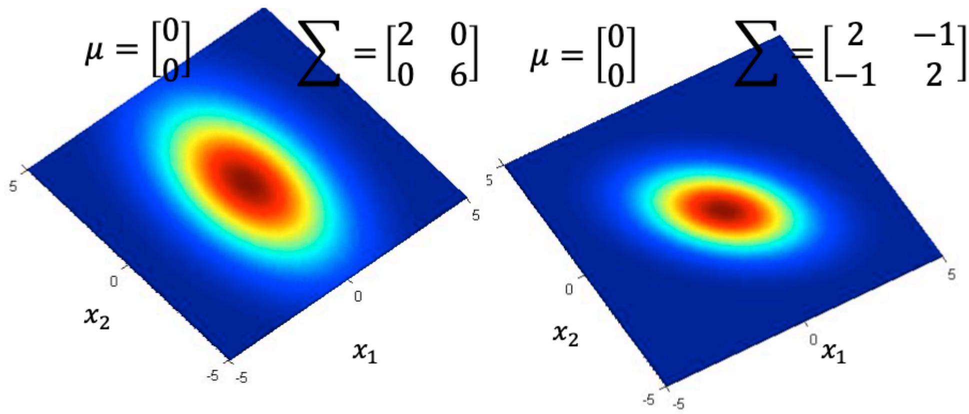
Gaussian Distribution

假设图中的点服从高斯分布，由于只考虑了两个属性，μ为一个二维向量，则有高斯分布公式，

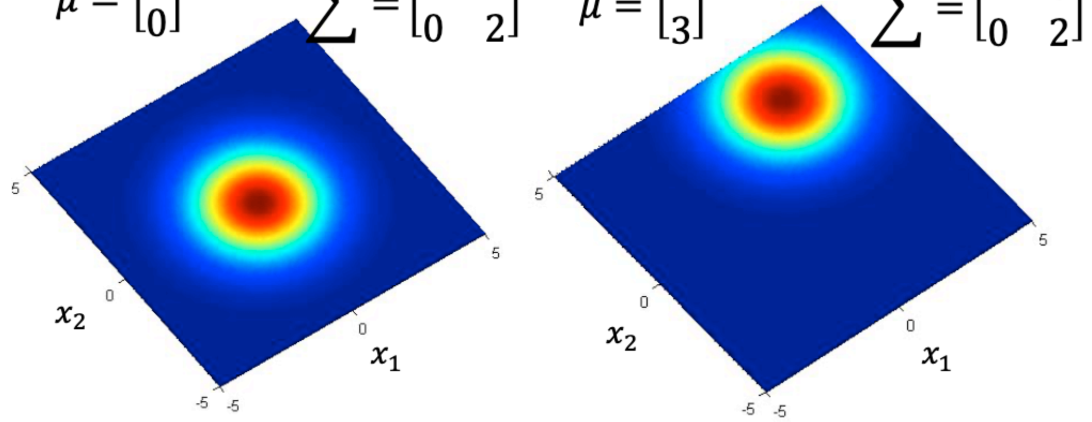
$$f_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

Input: vector x, output: probability of sampling x

The shape of the function determines by **mean $\boldsymbol{\mu}$** and **covariance matrix $\boldsymbol{\Sigma}$**



$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\sum = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ $\mu = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $\sum = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$



Maximum Likelihood

第i个example的概率密度函数为$f_{\mu,\sum}(x^i)$，可得出最大似然函数$L(\mu,\sum)$的表达式，其中$\mu=\mu^*,\sum=\sum^*$时，函数L得到最大值

We have the “Water” type Pokémons: $x^1, x^2, x^3, \ldots\ldots, x^{79}$

We assume $x^1, x^2, x^3, \ldots\ldots, x^{79}$ generate from the Gaussian (μ^*, Σ^*) with the ***maximum likelihood***

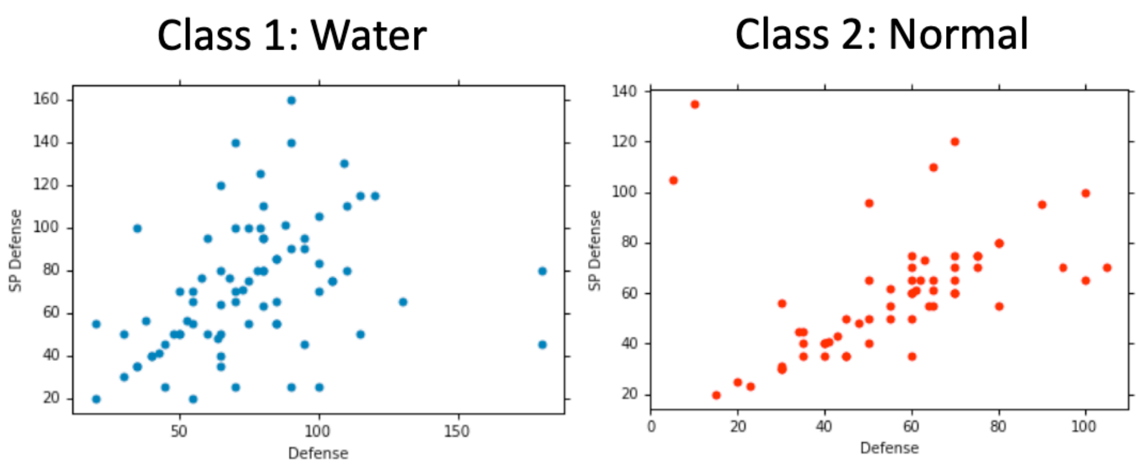
$$L(\mu,\Sigma) = f_{\mu,\Sigma}(x^1) f_{\mu,\Sigma}(x^2) f_{\mu,\Sigma}(x^3) \ldots\ldots f_{\mu,\Sigma}(x^{79})$$

$$f_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right\}$$

$$\mu^*, \Sigma^* = arg \max_{\mu,\Sigma} L(\mu,\Sigma)$$

$$\mu^* = \underbrace{\frac{1}{79}\sum_{n=1}^{79} x^n}_{\text{average}} \qquad \Sigma^* = \frac{1}{79}\sum_{n=1}^{79}(x^n-\mu^*)(x^n-\mu^*)^T$$

这里我们假设有两个类别，其参数和分布如下，

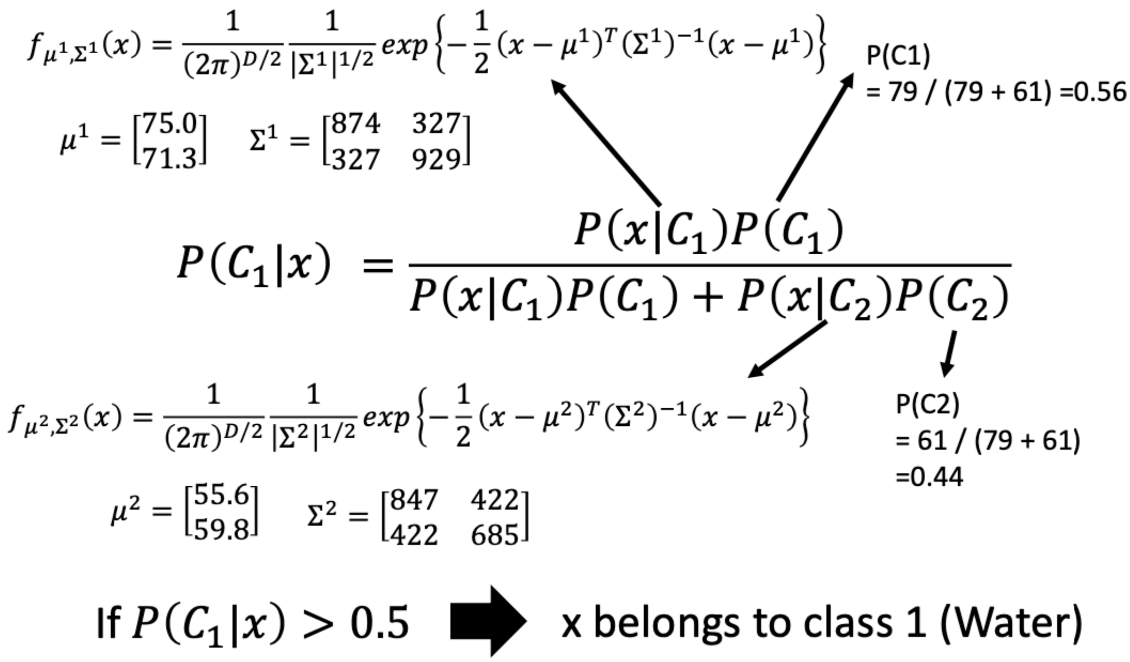


$$\mu^1 = \begin{bmatrix} 75.0 \\ 71.3 \end{bmatrix} \quad \Sigma^1 = \begin{bmatrix} 874 & 327 \\ 327 & 929 \end{bmatrix} \qquad \mu^2 = \begin{bmatrix} 55.6 \\ 59.8 \end{bmatrix} \quad \Sigma^2 = \begin{bmatrix} 847 & 422 \\ 422 & 685 \end{bmatrix}$$

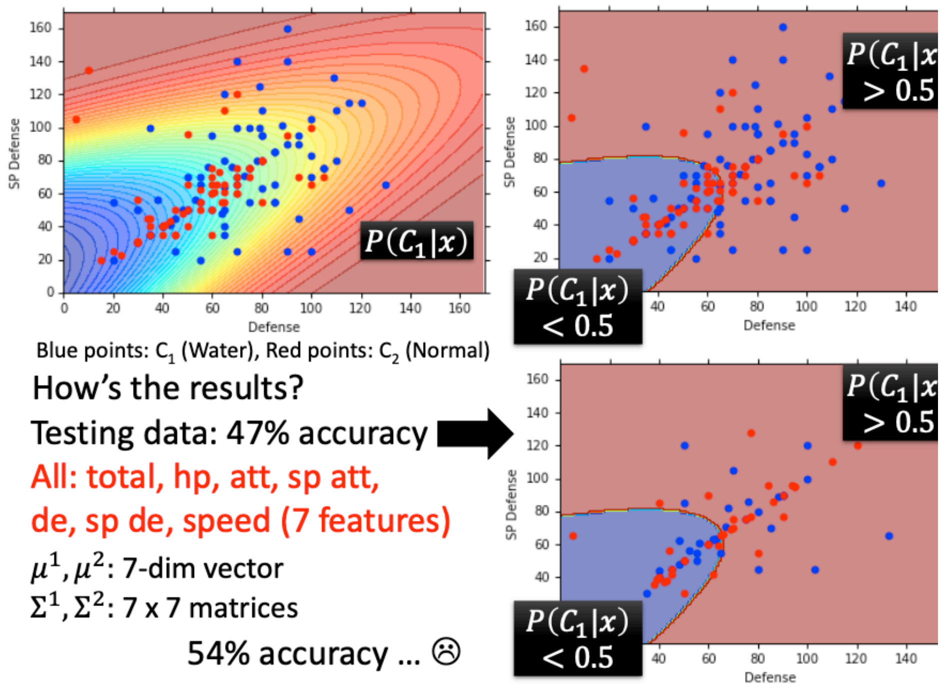
输入样例x是Class 1:Water的概率为，

$$P(C_1|x) = \frac{P(x|C_1)P(C_1)}{P(x)} = \frac{P(x|C_1)P(C_1)}{P(x|C_1)P(C_1) + P(x|C_2)P(C_2)}$$

再代入相应的表达式，其中$P(x|C_1) = f_{\mu^1, \sum^1}, P(x|C_2) = f_{\mu^2, \sum^2}$，即可算出x为C1的概率，如果算出这个概率大于0.5，我们就可以认为x的属性为C1

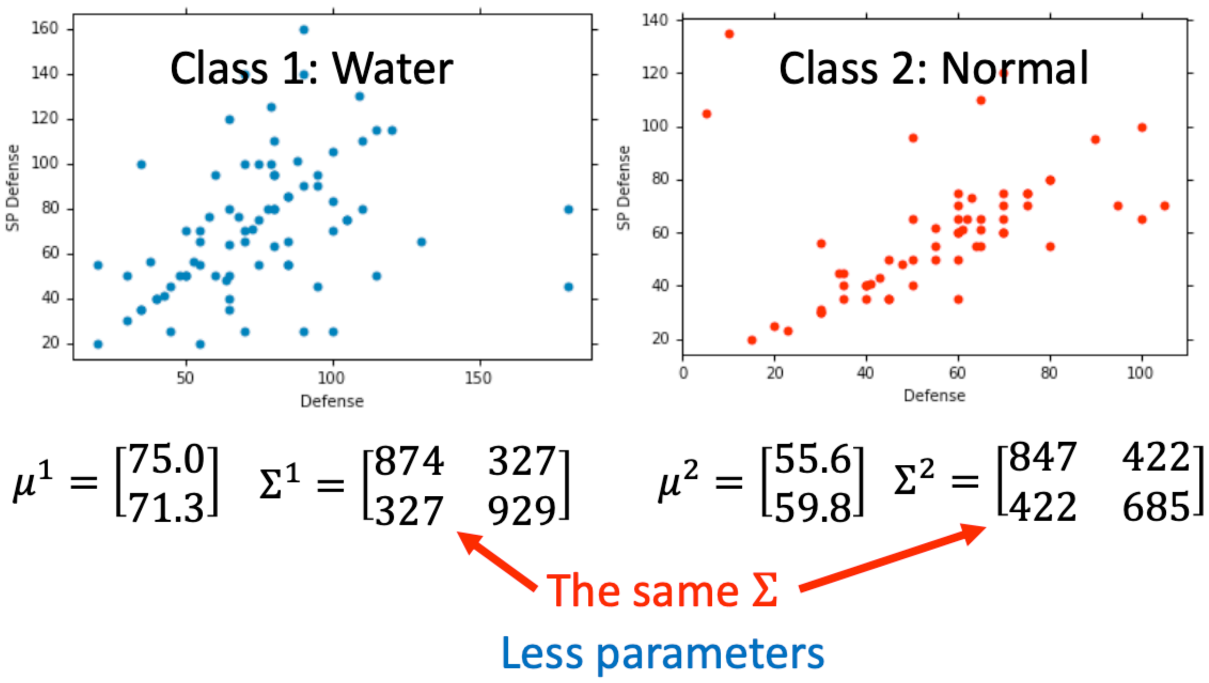


但这样子算出来的分类精确度很低，只有47%，就算加入其他的属性，精确度也只提高到了54%

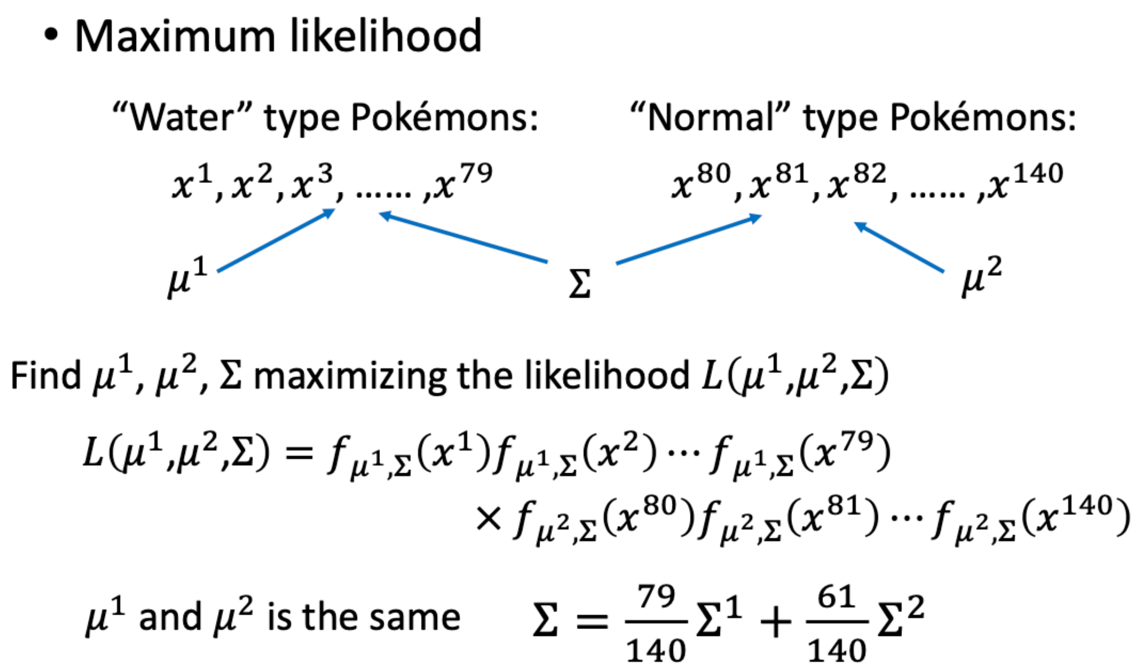


Modifying Model

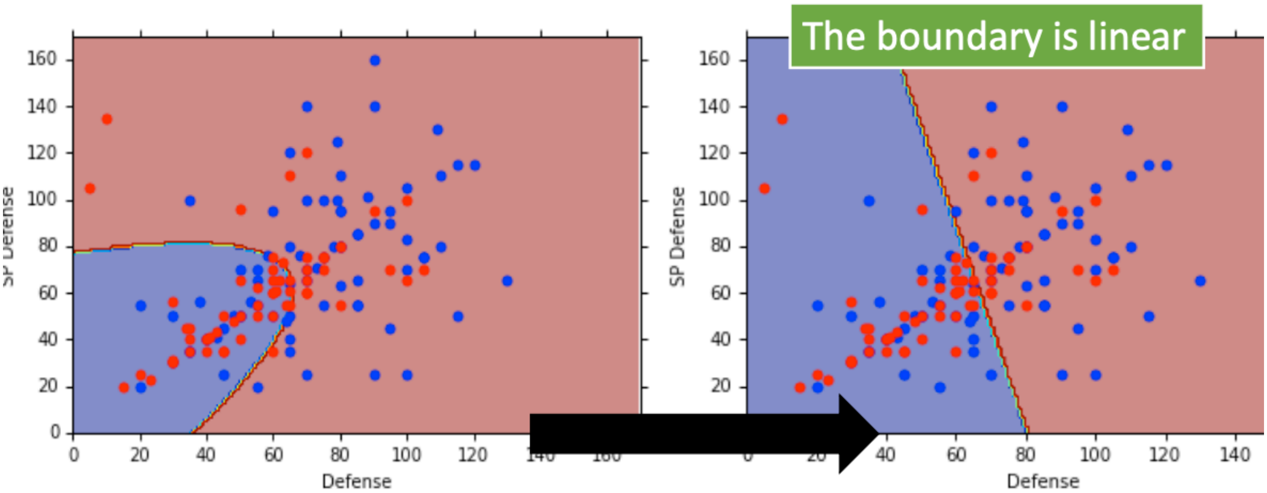
由于上面模型的精确度都不高，所以在此我们对模型进行了修改，模型的参数就只有三个，$\mu^1, \mu^2, \sum = \sum^1 = \sum^2$



对于Water属性的宝可梦对应参数为μ^1，Normal属性的宝可梦为μ^2，共同属性为$\sum$，这时对应的最大似然函数为$L(\mu^1, \mu^2, \sum)$，



修改后的模型准确率可以达到54%，如果加入更多的属性，准确率可以提高到73%



The same covariance matrix

All: total, hp, att, sp att, de, sp de, speed

54% accuracy → 73% accuracy

Three Steps

这里再回忆一下三个步骤：

（1）Function Set，计算分类为该类的概率$P(C_1|x)$，如果大于0.5，则认为类别为1，否则为类别2

（2）找出相对应的$\mu, \sum$，使得似然函数L取得最大值；

（3）得出使似然函数最大化的参数，$\mu^*, \sum^*$

$$u^* = \frac{1}{79}\sum_{n=1}^{79} x^n, \quad \sum^* = \frac{1}{79}\sum_{n=1}^{79}(x^n - \mu^*)(x^n - \mu^*)^T$$

- Function Set (Model):

x →

$$P(C_1|x) = \frac{P(x|C_1)P(C_1)}{P(x|C_1)P(C_1) + P(x|C_2)P(C_2)}$$

If $P(C_1|x) > 0.5$, output: class 1

Otherwise, output: class 2

- Goodness of a function:
  - The mean μ and covariance Σ that maximizing the likelihood (the probability of generating data)
- Find the best function: easy

Probability Distribution

这里我们所使用的概率分布是

$$P(x|C_1) = P(x_1|C_1)\ P(x_2|C_1)\ \cdots\cdots\ P(x_k|C_1)\ \cdots\cdots$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_K \end{bmatrix}$$

1-D Gaussian

For binary features, you may assume they are from Bernoulli distributions.

If you assume all the dimensions are independent, then you are using *Naive Bayes Classifier*.

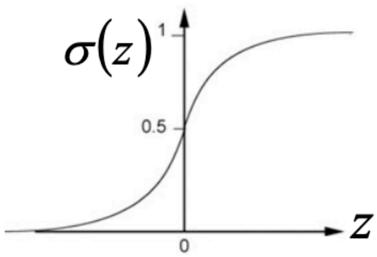
下面开始公式推导，

$$P(C_1|x) = \frac{P(x|C_1)P(C_1)}{P(x|C_1)P(C_1) + P(x|C_2)P(C_2)}$$

$$= \frac{1}{1 + \frac{P(x|C_2)P(C_2)}{P(x|C_1)P(C_1)}} = \frac{1}{1 + exp(-z)} = \sigma(z)$$

Sigmoid function

$$z = ln\frac{P(x|C_1)P(C_1)}{P(x|C_2)P(C_2)}$$



得出了我们的Sigmoid函数，$\sigma(z)$的函数图像为s型，值域范围为[0,1]，将z化简，并代入$P(C_i) = \frac{N_i}{N_1+N_2}$ (i=1,2)，

$$z = ln\frac{P(x|C_1)}{P(x|C_2)} + ln\boxed{\frac{P(C_1)}{P(C_2)}} \Rightarrow \frac{\frac{N_1}{N_1+N_2}}{\frac{N_2}{N_1+N_2}} = \frac{N_1}{N_2}$$

$$P(x|C_1) = \frac{1}{(2\pi)^{D/2}}\frac{1}{|\Sigma^1|^{1/2}}exp\left\{-\frac{1}{2}(x-\mu^1)^T(\Sigma^1)^{-1}(x-\mu^1)\right\}$$

$$P(x|C_2) = \frac{1}{(2\pi)^{D/2}}\frac{1}{|\Sigma^2|^{1/2}}exp\left\{-\frac{1}{2}(x-\mu^2)^T(\Sigma^2)^{-1}(x-\mu^2)\right\}$$

将$P(x|C_1), P(x|C_2)$的表达式代入$ln\frac{P(x|C_1)}{P(x|C_2)}$，

$$
ln\frac{\cancel{\frac{1}{(2\pi)^{D/2}}}\frac{1}{|\Sigma^1|^{1/2}}exp\left\{-\frac{1}{2}(x-\mu^1)^T(\Sigma^1)^{-1}(x-\mu^1)\right\}}{\cancel{\frac{1}{(2\pi)^{D/2}}}\frac{1}{|\Sigma^2|^{1/2}}exp\left\{-\frac{1}{2}(x-\mu^2)^T(\Sigma^2)^{-1}(x-\mu^2)\right\}}
$$

$$
=ln\frac{|\Sigma^2|^{1/2}}{|\Sigma^1|^{1/2}}exp\left\{-\frac{1}{2}[(x-\mu^1)^T(\Sigma^1)^{-1}(x-\mu^1) - (x-\mu^2)^T(\Sigma^2)^{-1}(x-\mu^2)]\right\}
$$

$$
=ln\frac{|\Sigma^2|^{1/2}}{|\Sigma^1|^{1/2}}-\frac{1}{2}[(x-\mu^1)^T(\Sigma^1)^{-1}(x-\mu^1)-(x-\mu^2)^T(\Sigma^2)^{-1}(x-\mu^2)]
$$

$$
\begin{aligned}
&(x-\mu^1)^T(\Sigma^1)^{-1}(x-\mu^1)\\
&=x^T(\Sigma^1)^{-1}x\underline{-x^T(\Sigma^1)^{-1}\mu^1-(\mu^1)^T(\Sigma^1)^{-1}x}+(\mu^1)^T(\Sigma^1)^{-1}\mu^1\\
&=x^T(\Sigma^1)^{-1}x\underline{-2(\mu^1)^T(\Sigma^1)^{-1}x}+(\mu^1)^T(\Sigma^1)^{-1}\mu^1\\
&(x-\mu^2)^T(\Sigma^2)^{-1}(x-\mu^2)\\
&=x^T(\Sigma^2)^{-1}x-2(\mu^2)^T(\Sigma^2)^{-1}x+(\mu^2)^T(\Sigma^2)^{-1}\mu^2
\end{aligned}
$$

$$
\boxed{\begin{aligned}
z=ln\frac{|\Sigma^2|^{1/2}}{|\Sigma^1|^{1/2}}&-\frac{1}{2}x^T(\Sigma^1)^{-1}x+(\mu^1)^T(\Sigma^1)^{-1}x-\frac{1}{2}(\mu^1)^T(\Sigma^1)^{-1}\mu^1\\
&+\frac{1}{2}x^T(\Sigma^2)^{-1}x-(\mu^2)^T(\Sigma^2)^{-1}x+\frac{1}{2}(\mu^2)^T(\Sigma^2)^{-1}\mu^2+ln\frac{N_1}{N_2}
\end{aligned}}
$$

再进行进一步化简，带入$\sum^1=\sum^2=\sum$，我们可以得出z的简易表达式$z=w\cdot x+b$，可得出$P(C_1|x)=\sigma(z)=\sigma(w\cdot x+b)$。当得出$N_1,N_2,\mu^1,\mu^2,\sum$时，就可以计算出w和b的值。

$$
\boxed{\begin{aligned} z = {} & \cancel{ln\frac{|\Sigma^2|^{1/2}}{|\Sigma^1|^{1/2}}} \cancel{-\frac{1}{2}x^T(\Sigma^1)^{-1}x} + (\mu^1)^T(\Sigma^1)^{-1}x - \frac{1}{2}(\mu^1)^T(\Sigma^1)^{-1}\mu^1 \\ & \cancel{+\frac{1}{2}x^T(\Sigma^2)^{-1}x} - (\mu^2)^T(\Sigma^2)^{-1}x + \frac{1}{2}(\mu^2)^T(\Sigma^2)^{-1}\mu^2 + ln\frac{N_1}{N_2} \end{aligned}}
$$

$$\Sigma_1 = \Sigma_2 = \Sigma$$

$$z = \underbrace{(\mu^1 - \mu^2)^T\Sigma^{-1}}_{\boldsymbol{w}^T}x \underbrace{- \frac{1}{2}(\mu^1)^T\Sigma^{-1}\mu^1 + \frac{1}{2}(\mu^2)^T\Sigma^{-1}\mu^2 + ln\frac{N_1}{N_2}}_{\text{b}}$$

$P(C_1|x) = \sigma(w \cdot x + b)$ How about directly find ***w*** and b?

In generative model, we estimate $N_1, N_2, \mu^1, \mu^2, \Sigma$

Then we have ***w*** and b