```
设有一对服从多元正态分布的变量(x,y)(\boldsymbol{x},\boldsymbol{y})(x,y),可以写出他们的联合概率密度函数:
 p\left(x,y\right)=N\left(\left[\mu x\mu y\right],\left[\Sigma xx\Sigma xy\Sigma yx\Sigma yy\right]\right)p(boldsymbol\{x\},boldsymbol\{y\})=b(heff(beff(befinal xay),boldsymbol\{y\})=b(heff(beff(befinal xay),boldsymbol\{y\})=b(heff(befinal xay),boldsymbol\{y\}
   \boldsymbol{Sigma}_{x y} \boldsymbol{Sigma}_{y x} \& \boldsymbol{Sigma}_{y y}\end{array} right]
p(x, y) = N ([\mu x \mu y], [\Sigma x x \Sigma y x \Sigma x y \Sigma y y])
 其中, \Sigma y x = \Sigma x y T \boldsymbol{\Sigma}_{y x}=\boldsymbol{\Sigma}_{x y}^{\mathrm{T}} \Sigma y x = \Sigma x y T.
   由舒尔补有:
 \label{thm:cc} $$ \cc} \boldsymbol{Sigma}_{x y} \boldsymbol{Sigma}_{y x} & \boldsymbol{Sigma}_{y x} \end{Sigma}_{y x} 
  y \end {array} 
 \mathcal{S}_{1}\end{ array}\right(x )\end{ array}\right(x )\end{ array}\end{ boldsymbol}\Sigma} {x x}-\boldsymbol}\Sigma} {x y}\boldsymbol}\Sigma} {y y}^{-1}
 \boldsymbol{Sigma}_{y x} \& \mathbb{0} \boldsymbol{Sigma}_{y y}\end \{array\} \mathbb{C} \mathbf{0} \& \boldsymbol{Sigma}_{y y}\end \{array\} \mathbb{C} \mathbf{0} \& \mathbf{0} \& \mathbf{0} \mathbf{
 \mathcal{Y} \ wrathbf{0} \ wrathbf{1}\end{array}\right]
 [\Sigma xx\Sigma yx\Sigma xy\Sigma yy] = [10\Sigma xy\Sigma yy-11][\Sigma xx - \Sigma xy\Sigma yy-1\Sigma yx00\Sigma yy][1\Sigma yy-1\Sigma yx01]
 对两边同时求逆有:
[\Sigma x x \Sigma x y \Sigma y x \Sigma y y] - 1 = [10 - \Sigma y y - 1 \Sigma y x 1][(\Sigma x x - \Sigma x y \Sigma y y - 1 \Sigma y x) - 100 \Sigma y y - 1][1 - \Sigma x y \Sigma y y - 101]
 y^{-1} \boldsymbol{(Sigma}_{y x}\right)^{-1} & \boldsymbol{(0)} \boldsymbol{(Sigma)_{y y}^{-1}} & \boldsymbol{(0)} \boldsymbol{(0)} & \boldsymbol{(
  1 \end {array} \left( x y \right) \left( x y \right
 \label{eq:linear_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_con
 因此,联合概率密度函数 p(x,y) p(boldsymbol\{x\}, boldsymbol\{y\}) p(x,y) 指数部分的二次项为:
```

 $\label{leff(begin{array} {cc} \boldsymbol{\Sigma}_{x x} \& \boldsymbol{\Sigma}_{y x} & \boldsymbol{\Sigm$ $y \end {array} \end {array}$ & $\mbox{mathb} f{1}\end{array} \left[\mbox{xy} \left(\mbox{xy} \right) \left(\mbox{xy} \right) \right] \left(\mbox{xy} \mbox{ymbol} \right) \left(\mbox{xy} \mbox{ymbox{ymbol} \right) \left(\mbox{xy} \mbox{ymbox{ymbol} \right) \left(\mbox{xy} \mbox{y$

```
xy\Sigma yy - 1\Sigma yx) - 100\Sigma yy - 1] \times [1 - \Sigma xy\Sigma yy - 101]([xy] - [\mu x\mu y]) = (x - \mu x - \Sigma xy\Sigma yy - 1(y - \mu y))T(\Sigma xx - \mu x - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy - 1(y - \mu y))T(\Sigma xy - \mu xy\Sigma yy 
\sum x y \sum y y - 1 \sum y x) - 1 \times (x - \mu x - \sum x y \sum y y - 1 (y - \mu y)) + (y - \mu y) T \sum y y - 1 (y - \mu y) \setminus (y - \mu y)
 {I}\boldsymbol{x} \\\boldsymbol{y}\end{array}\right]-\left[\begin{array} {I}\boldsymbol{\mu} {x}
\end{array} $$\{l\} \boldsymbol{x} \boldsymbol{y}\end{array} \right]-\left[\begin{array} {l}\boldsymbol{\mu}_{x} \end{array} \right]-\left[\begin{array} {l}\boldsymbol{\mu}_{x} \end{array} \right]-\end{array} $$\{l\}\boldsymbol{\mu}_{x} \end{array} $$\{l\}\boldsymbol{\mu}_{x} \end{array} $$\{l\}\boldsymbol{\mu}_{x} \end{array} $$\{l\}\boldsymbol{\mu}_{x} \end{array} \right]-\end{array} $$\{l\}\boldsymbol{\mu}_{x} \end{array} $$\{l\}\boldsymbol{\mu}
\label{thm:continual} $$\left( y\right) \left( x^{T} \right) \left( x^{T} 
\boldsymbol{Sigma}_{y y}^{-1} \boldsymbol{Sigma}_{y x} \& \boldsymbol{1}\end {array} \right] \label{Sigma}_{y x} \\
 \label{linear_xy} $$ \left( \sum_{x \in \mathbb{S}_{x \in \mathbb{S
\boldsymbol{0} $$\boldsymbol{Sigma}_{yy}^{-1}\end{array}\right] \& \times\left[\begin{array}{cc}\right] \& -1 \end{array}\right] $$
\boldsymbol{Sigma} \{x y\} \boldsymbol{Sigma} \{y y\}^{-1} \label{final} \& \boldsymbol{Sigma} \ array\\ \ array\\ \boldsymbol{Sigma} \ array\\ \bol
 {I}\boldsymbol{x} \\\boldsymbol{y}\end {array}\right]-\left[\begin{array} {I}\boldsymbol{\mu} {x}
\boldsymbol{Sigma}_{y}^{-1}\left[boldsymbol{y}-boldsymbol{mu}_{y}\cdot (y)^{-1}\left[boldsymbol{Sigma}_{x}-(y)\right].
\boldsymbol{\Sigma}_{x y} \boldsymbol{\Sigma}_{y y}^{-1} \boldsymbol{\Sigma}_{y x}\right]^{-1} \boldsymbol{\Sigma}_{x y} \boldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymboldsymbold
\label{thmu} $$ \boldsymbol{\sigma}_{x y} \boldsymbol{\sigma}_{y y}^{-1}\left(boldsymbol{y}-boldsymbol{\sigma}_{y y}^{-1}\right). $$
\boldsymbol\{mu\}_{y} \right) + \left[ (boldsymbol\{y\}-boldsymbol\{mu\}_{y} \right] + \left[ (boldsymbol\{y\}-boldsymbol\{mu\}_{y} \right] + \left[ (boldsymbol\{Ngma\}_{y} \right] + \left[ (boldsymbol\{mu\}_{y} \right] + \left
 1}\left(\boldsymbol{y}-\boldsymbol{\mu}_{y}\right)\end{aligned}
 ==([xy]-[\mu x\mu y])T[\Sigma xx\Sigma yx\Sigma xy\Sigma yy]-1([xy]-[\mu x\mu y])([xy]-[\mu x\mu y])T[1-\Sigma yy-1\Sigma yx01][(\Sigma xx-\Sigma xy\Sigma yy-1\Sigma yx)-100\Sigma yy-1]\times [10-\Sigma xy\Sigma yy-11]
 ([xy] - [\mu x \mu y])(x - \mu x - \Sigma xy \Sigma yy - 1(y - \mu y))T(\Sigma xx - \Sigma xy \Sigma yy - 1\Sigma yx) - 1 \times (x - \mu x - \Sigma xy \Sigma yy - 1(y - \mu y)) + (y - \mu y)T\Sigma yy - 1(y - \mu y)
很明显可以看出,这是两个二次项的和。
```

又由贝叶斯公式有:

 $p(x, y) = p(x \mid y) p(y) p(boldsymbol\{x\}, boldsymbol\{y\}) = p(boldsymbol\{x\}, boldsymbol\{y\}) p(boldsymbol\{y\}) p(boldsymbol\{y\}, boldsymbol\{y\})$ $p(x, y) = p(x \mid y)p(y)$

并且:

 $p\left(y\right) = N\left(\mu y, \Sigma yy\right) \\ p\left(\text{boldsymbol}\{N\} \setminus \{y\}, \text{boldsymbol}\{Sigma\}_{y} \setminus \{y\} \setminus \{y\} \setminus \{y\}, \text{boldsymbol}\{Sigma\}_{y} \setminus \{y\} \setminus \{y\}, \text{boldsymbol}\{Sigma\}_{y} \setminus \{y\} \setminus \{y\}, \text{boldsymbol}\{Sigma\}_{y} \setminus \{y\}, \text{boldsymbol}\{S$ 因此,由幂运算中同底数幂相乘,底数不变、指数相加的性质,可以得到:

 $p(x \mid y) = N(\mu x + \Sigma x y \Sigma y y - 1(y - \mu y), \Sigma x x - \Sigma x y \Sigma y y - 1\Sigma y x) p(\text{boldsymbol}\{x\} \text{ mid boldsymbol}\{y\})$

 $$$ = \mathcal{N}\left(\frac{x} + \boldsymbol{x} + \boldsymbol{x} - \boldsymbol{x} \right) \left(\frac{x y} - \frac{y}^{-1} \left(\frac{y y}^{-1} \right) \left(\frac{y y}{-1} \right) \left(\frac{y y}{-1}$

这便是高斯推断中最重要的部分: 从状态的先验概率分布出发, 然后基于一些观测值来缩小这个范围。