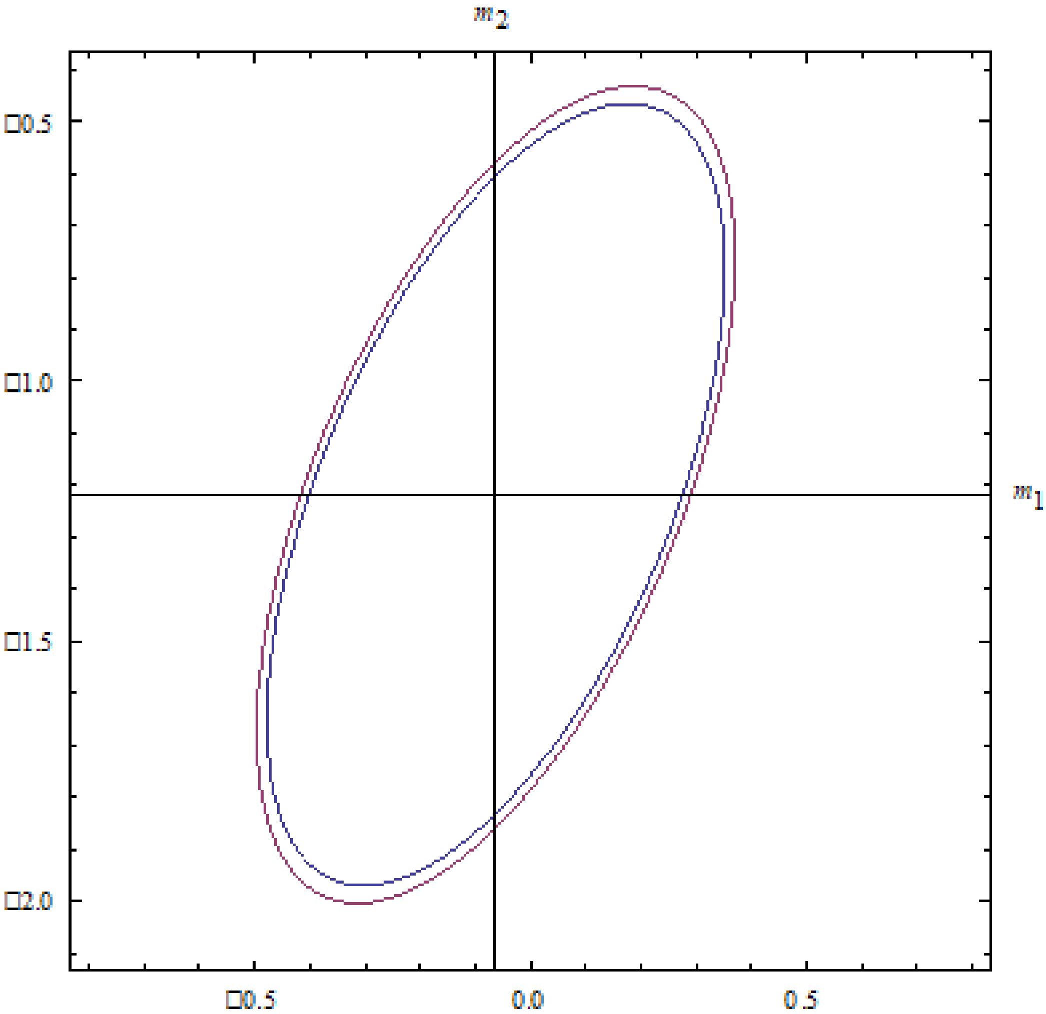

Рис. 1: Доверительная область для математических жиданий двумерного случайного вектора при известной матрице коварияций и при неизвестной

2.3 Критерии для проверки гипотез о параметрах многомерной нормальной случайной величины на основе отношения правдоподобия

Гипотеза о равенстве матриц ковариаций

Пусть имеются выборки $X_1, X_2, \ldots, X_q$ объемов n_1, n_2, …,n_q соответственно из p-мерных нормальных совокупностей $N(\vec{\mu}_1, \Sigma_1)$, $N(\vec{\mu}_2, \Sigma_2)$, …, $N(\vec{\mu}_q, \Sigma_q)$. Проверяется гипотеза H_0 : $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_q$.

Отношение правдоподобия в этом случае будет равно:

$$\lambda = \frac{\prod\limits_{i=1}^{q} |S_i|^{\frac{n_i-1}{2}}}{|S|^{\frac{n-q}{2}}},$$

где $S = \frac{1}{n-q}\sum\limits_{i=1}^{q}(n_i - 1)S_i$, а S_i ОМП ковариационной матрицы для i-ой выборки.

Если положить: $\rho = 1 - \left(\sum\limits_{i=1}^{q} \frac{1}{n_i-1} - \frac{1}{n-q}\right) \frac{2p^2+3p-1}{6(p+1)(q-1)}$, то статистика:

$$\eta = -2\rho \ln \lambda = -\rho \left(\sum_{i=1}^{q} (n_i - 1) \ln |S_i| - (n-q) \ln |S| \right)$$

будет асимптотически иметь распределение χ^2 с $\nu = \frac{1}{2}(q-1)p(p+1)$ степенями свободы.

Таким образом, критерий будет иметь вид:

$$\delta(X, Y) = \begin{cases} H_0, & -2\rho \ln \lambda < \tau_{1-\alpha} \\ H_1, & -2\rho \ln \lambda \geq \tau_{1-\alpha} \end{cases},$$

где $\tau_{1-\alpha}$ - квантиль распределения χ^2 с $\nu = \frac{1}{2}p(p+1)$ степенями свободы.

Пример. По выборкам объемов $n_1 = 20$ и $n_2 = 14$ из двумерных нормальных совокупностей $N(\vec{\mu}_1, \Sigma_1)$ и $N(\vec{\mu}_2, \Sigma_2)$ соответственно получены выборочные матрицы ковариаций:
$S_1 = \begin{pmatrix} 1,50 & 0,86 \\ 0,86 & 1,31 \end{pmatrix}$, $S_2 = \begin{pmatrix} 2,92 & 3,08 \\ 3,08 & 4,05 \end{pmatrix}$.

Требуется на уровне значимости 0,05 проверить гипотезу $H_0 : \Sigma_1 = \Sigma_2$, против альтернативы $H_1 : \Sigma_1 \neq \Sigma_2$.

Находим обобщенную выборочную матрицу ковариаций:

$$S = \frac{1}{n_1 + n_2 - 2} \left[(n_1 - 1)S_1 + (n_2 - 1)S_2 \right] \approx \begin{pmatrix} 2,077 & 1,762 \\ 1,762 & 2,423 \end{pmatrix}.$$

Вычисляем поправочный множитель:

$$\rho = 1 - \left(\sum_{i=1}^{2} \frac{1}{n_i - 1} - \frac{1}{n-2} \right) \frac{2p^2 + 3p - 1}{6(p+1)} \approx 0,929$$

Вычисляем наблюдаемое значение статистики критерия:

$$\eta_{\text{выб.}} = -2\rho \ln \lambda = -\rho \left(\sum_{i=1}^{2} (n_i - 1) \ln |S_i| - (n-2) \ln |S| \right) \approx 5,668.$$

Находим критическое значение статистики – критическую точку уровня 0,05 распределения χ^2 с $\nu = \frac{1}{2}p(p+1) = 3$ степенями свободы: $\eta_{\text{крит.}} \approx 7,815$. Так как наблюдаемое значение статистики $\eta_{\text{выб.}} = 5,668$ меньше критического $\eta_{\text{крит.}} \approx 7,815$ гипотезу $H_0 : \Sigma_1 = \Sigma_2$ на уровне значимости 0,05 принимаем.

Гипотеза о равенстве векторов средних

Пусть имеются выборки $X^{(1)}, X^{(2)}, \ldots, X^{(q)}$ объемов $n_1, n_2, \ldots, n_q$, $\sum\limits_{i=1}^{q} n_i = n$ соответственно из p-мерных нормальных совокупностей $N(\vec{\mu}_1, \Sigma)$, $N(\vec{\mu}_2, \Sigma)$, …, $N(\vec{\mu}_q, \Sigma)$. Проверяется гипотеза $H_0 : \vec{\mu}_1 = \vec{\mu}_2 = \ldots = \vec{\mu}_q$.

Отношение правдоподобия:

$$\lambda = \frac{|S|^{\frac{n-q}{2}}}{|S_0|^{\frac{n-q}{2}}},$$

где:

$$S = \frac{1}{n-q}\sum_{i=1}^{q}(n_i - 1)S_i$$

$$S_0 = \frac{1}{n-q}\hat{X}^T\hat{X} = S + \frac{1}{n-q}\sum_{i=1}^{q} n_i(\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^T.$$

Здесь $\hat{X}$- центрированная матрица объединенной выборки, $\bar{X}$ - вектор выборочных средних объединенной выборки, $\bar{X}_i$ - вектор выборочных средних i-ой выборки.

Если положить

$$\rho = 1 - \frac{p-q+2}{2(n-q)},$$

то статистика $\eta = -2\rho\ln\lambda$ будет иметь распределение χ^2 с $\nu = p(q-1)$ степенями свободы.

Гипотеза однородности многомерных нормальных совокупностей.

Пусть имеются выборки $X^{(1)}, X^{(2)}, \ldots, X^{(q)}$ объемов $n_1,\ n_2,\ \ldots, n_q$ соответственно из p-мерных нормальных совокупностей $N(\vec{\mu}_1, \Sigma_1)$, $N(\vec{\mu}_2, \Sigma_2)$, …, $N(\vec{\mu}_q, \Sigma_q)$. Проверяется гипотеза $H_0 : \vec{\mu}_1 = \vec{\mu}_2 = \ldots = \vec{\mu}_q, \quad \Sigma_1 = \Sigma_2 = \ldots = \Sigma_q$.

Отношение правдоподобия для данной гипотезы:

$$\lambda = \frac{\prod\limits_{i=1}^{q}|S_i|^{\frac{n_i-1}{2}}}{|S_0|^{\frac{n-q}{2}}},$$

где

$$S = \frac{1}{n-q}\sum_{i=1}^{q}(n_i - 1)S_i$$

$$S_0 = \frac{1}{n-q}\hat{X}^T\hat{X} = S + \frac{1}{n-q}\sum_{i=1}^{q} n_i(\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^T.$$

Если положить (Т. Андерсон. Введение в многомерный статистический анализ):

$$\rho = 1 - \left(\sum_{i=1}^{q}\frac{1}{n_i - 1} - \frac{1}{n-q}\right)\frac{2p^2+3p-1}{6(p+3)(q-1)} - \frac{1}{n-q}\frac{p-q+2}{p+3},$$

то статистика $\eta = -2\rho\ln\lambda$ будет иметь распределение χ^2 с числом степеней свободы

$$\nu = \frac{1}{2}(q-1)p(p+3).$$