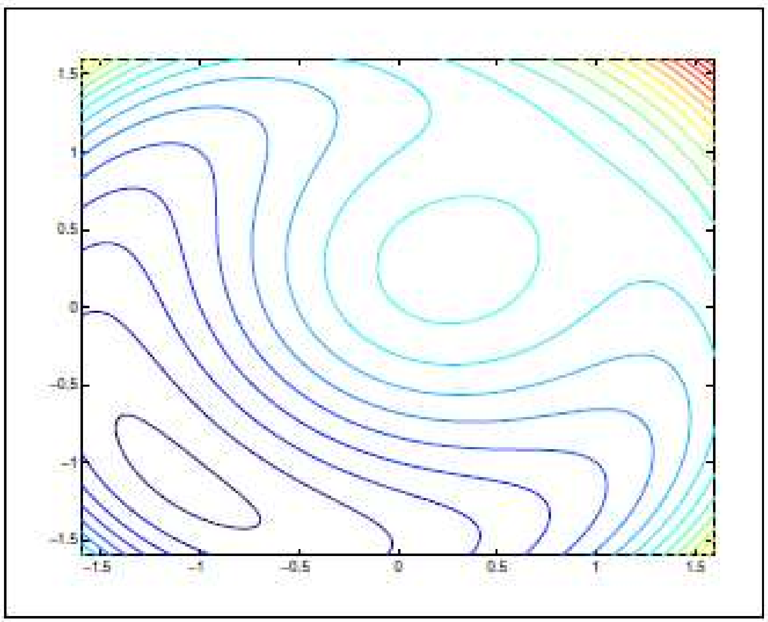
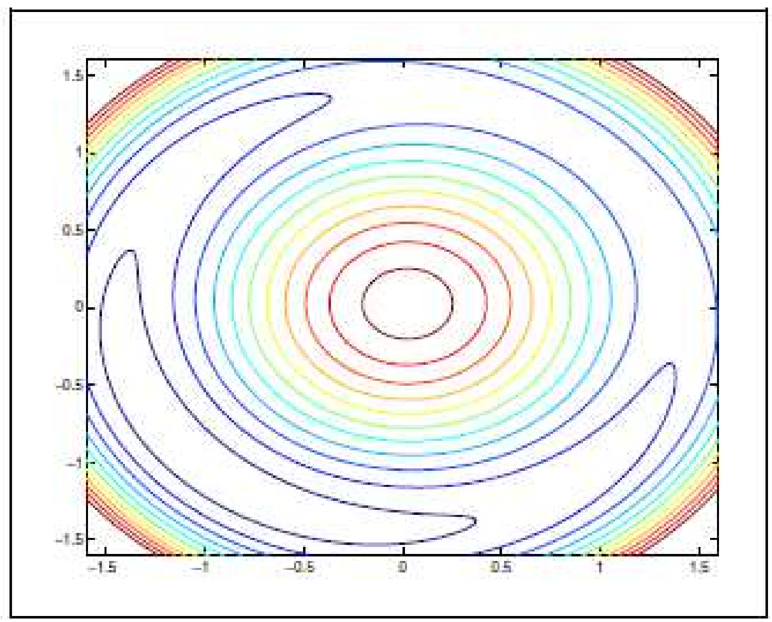
Quadratic penalty method

$$\min\ x_1 + x_2 \qquad x_1^2 + x_2^2 - 2 = 0,$$

the solution is $(-1, -1)^T$



μ=1



μ=10

Ill-conditioness of Hessian

$$\min_x (x_1 + x_2) \text{ s.t. } \mathrm{x}_1^2 + x_2^2 = 2$$

$$Q(x, \mu) = x_1 + x_2 + \frac{\mu}{2}(x_1^2 + x_2^2 - 2)^2$$

$$\nabla_{xx} Q = \begin{bmatrix} 2\mu(3x_1^2 + x_2^2 - 2) & 4\mu x_1 x_2 \\ 4\mu x_1 x_2 & 2\mu(x_1^2 + 3x_2^2 - 2) \end{bmatrix}$$

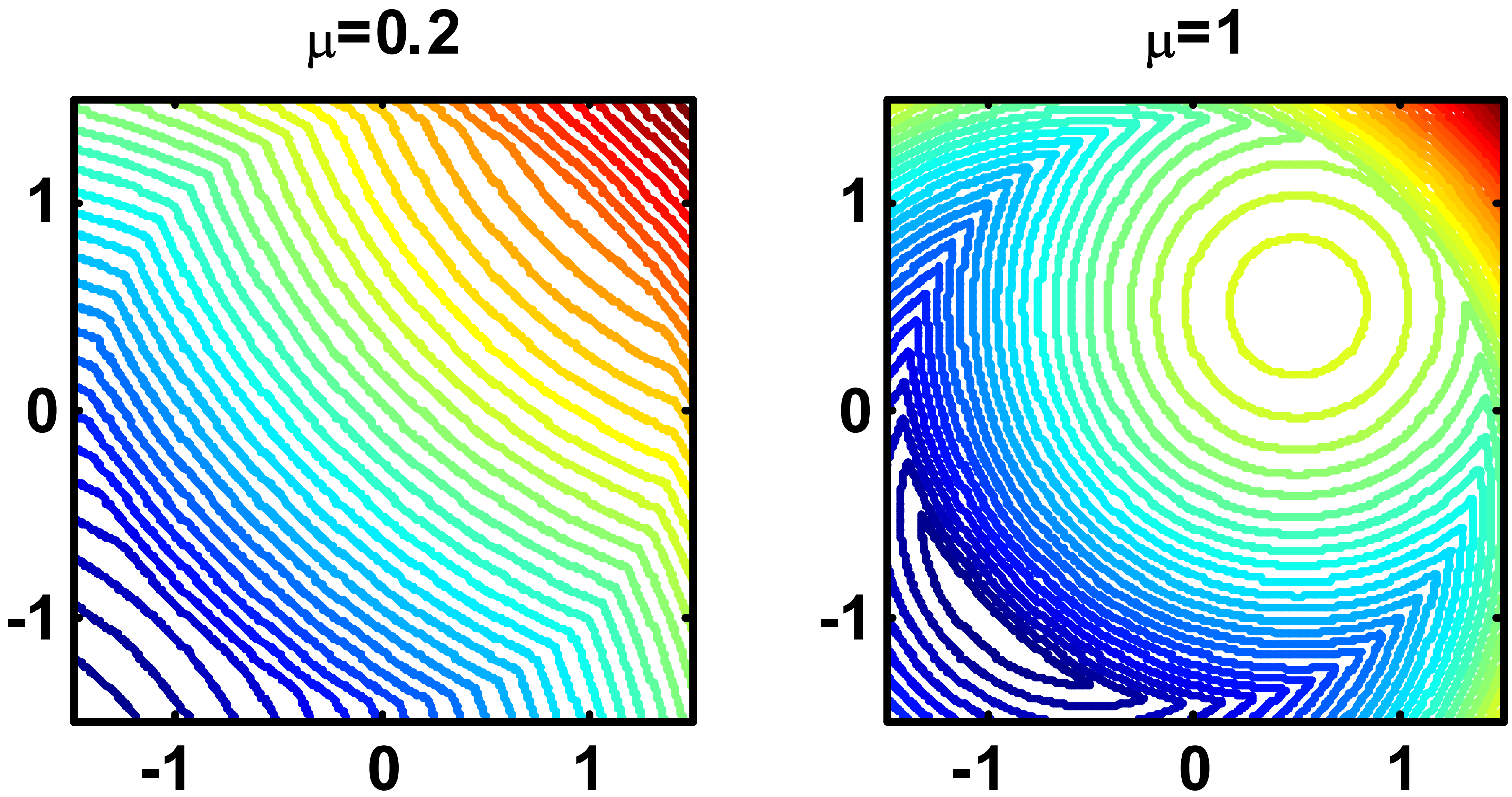
As it converges $x_1^2 + x_2^2 \to 2$

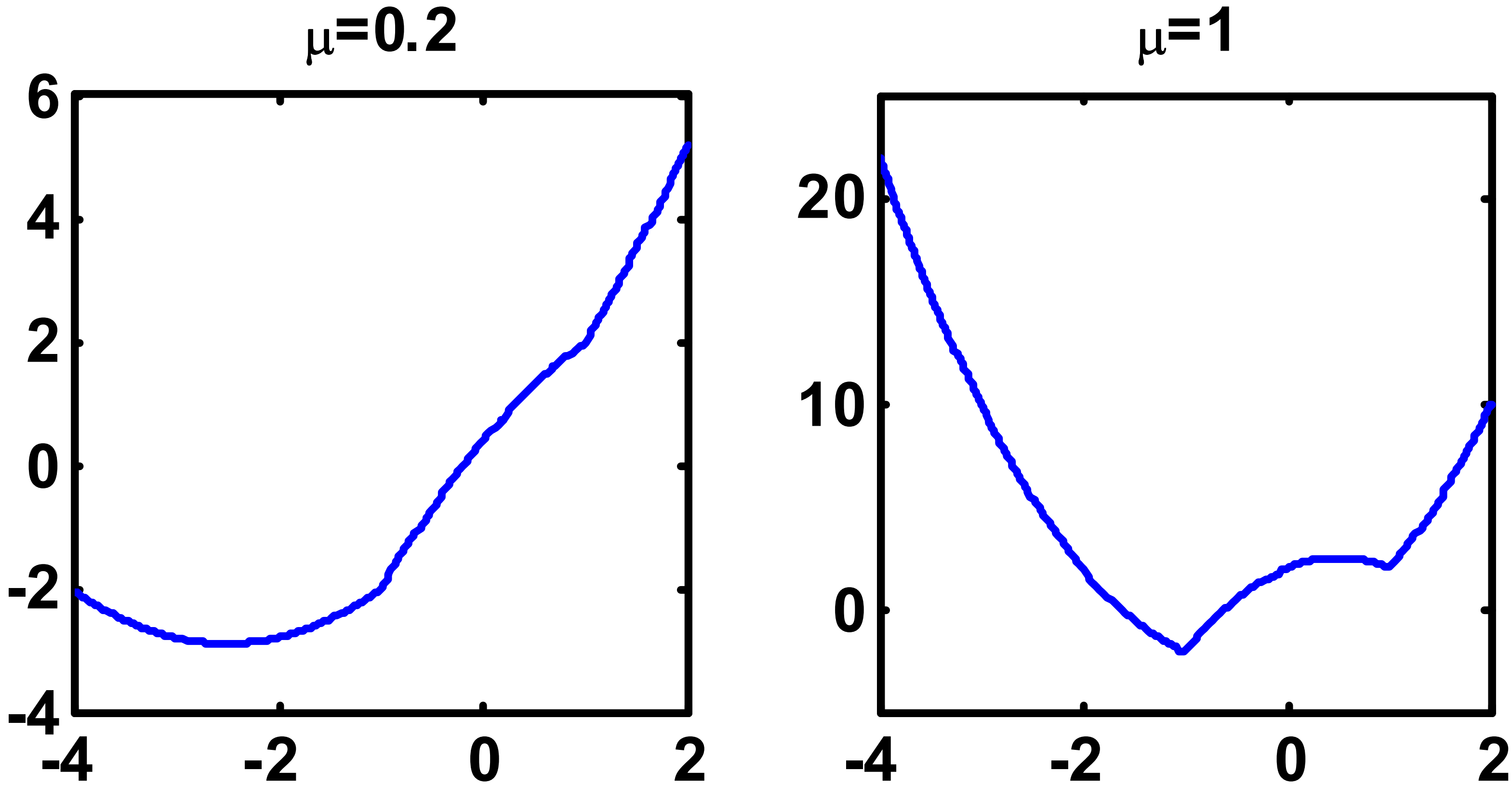
$$\nabla_{xx} Q \to \begin{bmatrix} 4\mu x_1^2 & 4\mu x_1 x_2 \\ 4\mu x_1 x_2 & 4\mu x_2^2 \end{bmatrix} = 4\mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

ℓ1 penalty method

$$\min_{x}(x_1 + x_2) \text{ s.t. } \mathrm{x}_1^2 + x_2^2 = 2$$

$$\phi_1(x, \mu) = x_1 + x_2 + \mu \left| x_1^2 + x_2^2 - 2 \right|$$





- The minimum μ that makes ϕ_1 exact is 0.5.

Augmented Lagrangian

$$L_A(x,\mu) = x_1 + x_2 - \lambda(x_1^2 + x_2^2 - 2) + \frac{\mu}{2}(x_1^2 + x_2^2 - 2)^2$$

$$\nabla_{xx} L_A = \begin{bmatrix} 2\mu(3x_1^2 + x_2^2 - 2) - 2\lambda & 4\mu x_1 x_2 \\ 4\mu x_1 x_2 & 2\mu(x_1^2 + 3x_2^2 - 2) - 2\lambda \end{bmatrix}$$

As it converges $x_1^2 + x_2^2 \to 2$

$$\nabla_{xx} L_A \to \begin{bmatrix} 4\mu x_1^2 - 2\lambda & 4\mu x_1 x_2 \\ 4\mu x_1 x_2 & 4\mu x_2^2 - 2\lambda \end{bmatrix}$$

To make the augmented Lagrangian exact, i.e. having the same solution as the original problem, what is the minimum μ*? (Theorem 17.5)

$$\nabla_{xx} L_A = \begin{bmatrix} 2\mu(3x_1^2 + x_2^2 - 2) - 2\lambda & 4\mu x_1 x_2 \\ 4\mu x_1 x_2 & 2\mu(x_1^2 + 3x_2^2 - 2) - 2\lambda \end{bmatrix}$$

Optimal solution of the original problem:
x* is (-1,-1), λ* is -½.

$$\nabla_{xx} L_A(x^*, \lambda^*, \mu) = \begin{bmatrix} 4\mu + 1 & 4\mu \\ 4\mu & 4\mu + 1 \end{bmatrix} (= H_A^*)$$

must be positive definite (2^{nd} order condition, chap 12). The eigenvalues of H^*_A are 1 and 8μ+1. The minimum $\mu^*=-1/8$.