Optimization Method
Solmaz Kia
Mechanical and Aerospace Eng. Dept.,
University of California Irvine

Lecture 14
Penalty Function Method

Consult: Chapter 12 of Ref[2] and Chapter 17 of Ref[3]

Solution methods for constrained optimization

- Idea: Seek the solution by replacing the original constrained problem by a sequence of unconstrained sub-problems
  - Penalty method
  - Barrier method
  - Augmented Lagrangian method

Quadratic Penalty Method

Motivation:

- the original objective of the constrained optimization problem, plus
- one additional term for each constraint, which is positive when the current point x violates that constraint and zero otherwise.

Most approaches define a sequence of such penalty functions, in which the penalty terms for the constraint violations are multiplied by a positive coefficient. By making this coefficient larger, we penalize constraint violations more severely, thereby forcing the minimizer of the penalty function closer to the feasible region for the constrained problem. The simplest penalty function of this type is the quadratic penalty function , in which the penalty terms are the squares of the constraint violations.

Penalty Method

$$\underset{x \in \Omega}{\text{Minimize}} \quad f(x) \quad \text{subject to} \qquad (1)$$

The idea of a penalty function method:

replace problem (1) by an unconstrained problem of the form

$$\text{Minimize} \quad f(x)+c\,P(x) \qquad (2)$$

where c is a positive constant (penalty weight) and P is a function on R^n satisfying: (i) P is continuous, (ii) $P(x) > 0$ for all $x \in R^n/\Omega$, and (iii) $P(x) = 0$ if and only if $x \in \Omega$.

Quadratic Penalty Method

$$\boxed{\begin{array}{ll} \text{minimize } f(x) & \text{subject to} \\ \mathrm{h_i}(x) = 0, & \text{i=1,...,m} \\ g_{\mathrm{j}}(x) \le 0, & \text{j=1,...,r} \end{array}}$$

$$\text{minimize } f(x) + \underbrace{\frac{\mathsf{c}}{2}\sum_{i=1}^{m} \mathrm{h_i}(x)^2 + \frac{\mathsf{c}}{2}\sum_{j=1}^{r}(\max\{0, g_j(x)\})^2}_{\text{c P(x)}}$$

Consider

$$\min x_1 + x_2 \quad \text{subject to } x_1^2 + x_2^2 - 2 = 0, \tag{1}$$

for which the solution is $(-1, -1)^T$ and the quadratic penalty function is

$$Q(x; c) = x_1 + x_2 + \frac{c}{2}\left(x_1^2 + x_2^2 - 2\right)^2. \tag{2}$$

We plot the contours of this function in Figures 2 and 3. In Figure 2 we have $c = 1$, and we observe a minimizer of Q near the point $(-1.1, -1.1)^T$. (There is also a local maximizer near $x = (0.3, 0.3)^T$.) In Figure 3 we have $c = 10$, so points that do not lie on the feasible circle defined by $x_1^2 + x_2^2 = 2$ suffer a much greater penalty than in the first figure—the "trough" of low values of Q is clearly evident. The minimizer in this figure is much closer to the solution $(-1, -1)^T$ of the problem (1). A local maximum lies near $(0, 0)^T$, and Q goes rapidly to ∞ outside the circle $x_1^2 + x_2^2 = 2$. ❐

$$\min x_1 + x_2 \quad \text{subject to } x_1^2 + x_2^2 - 2 = 0, \tag{1}$$

Figure 1

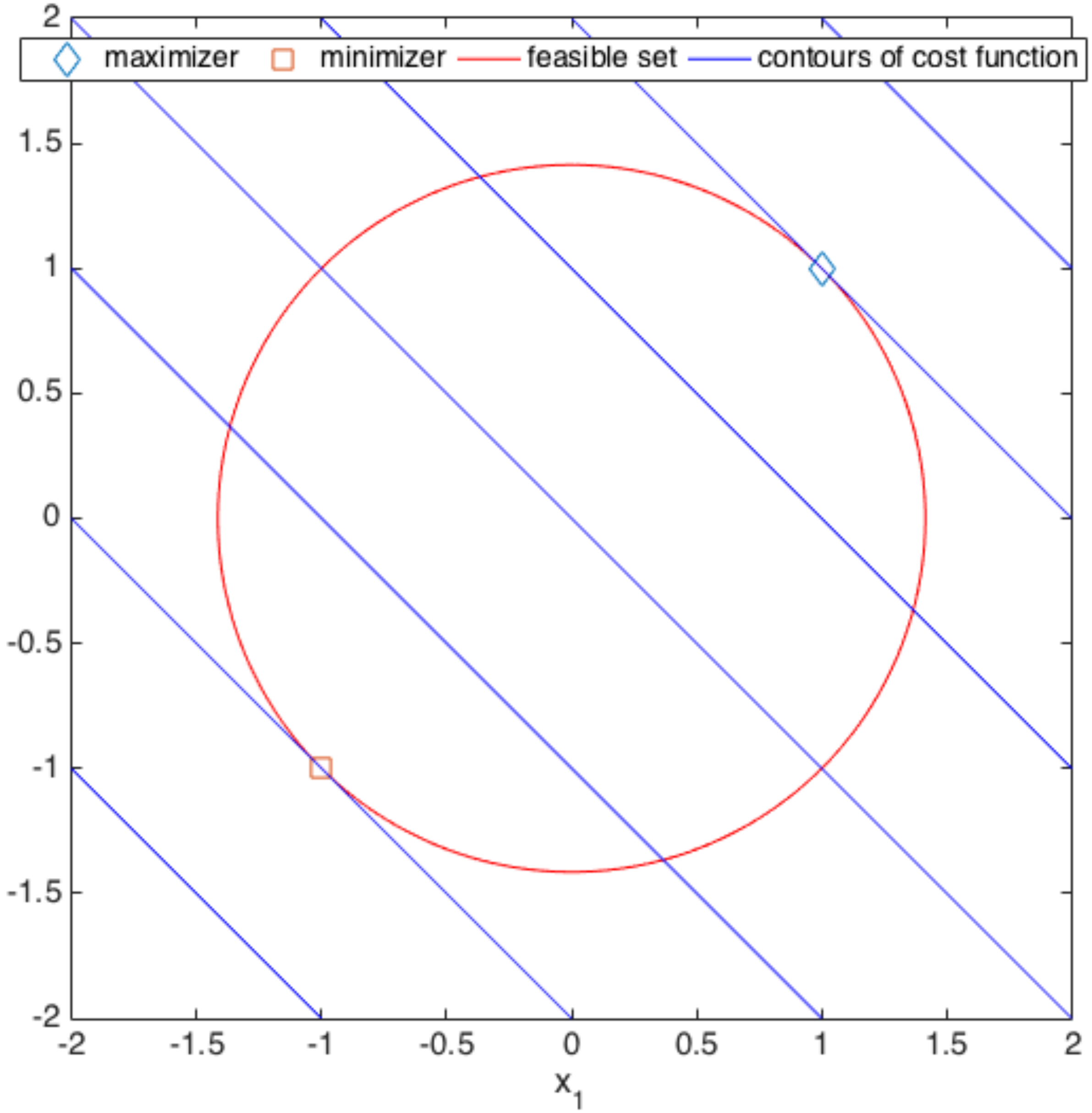


Figure 2

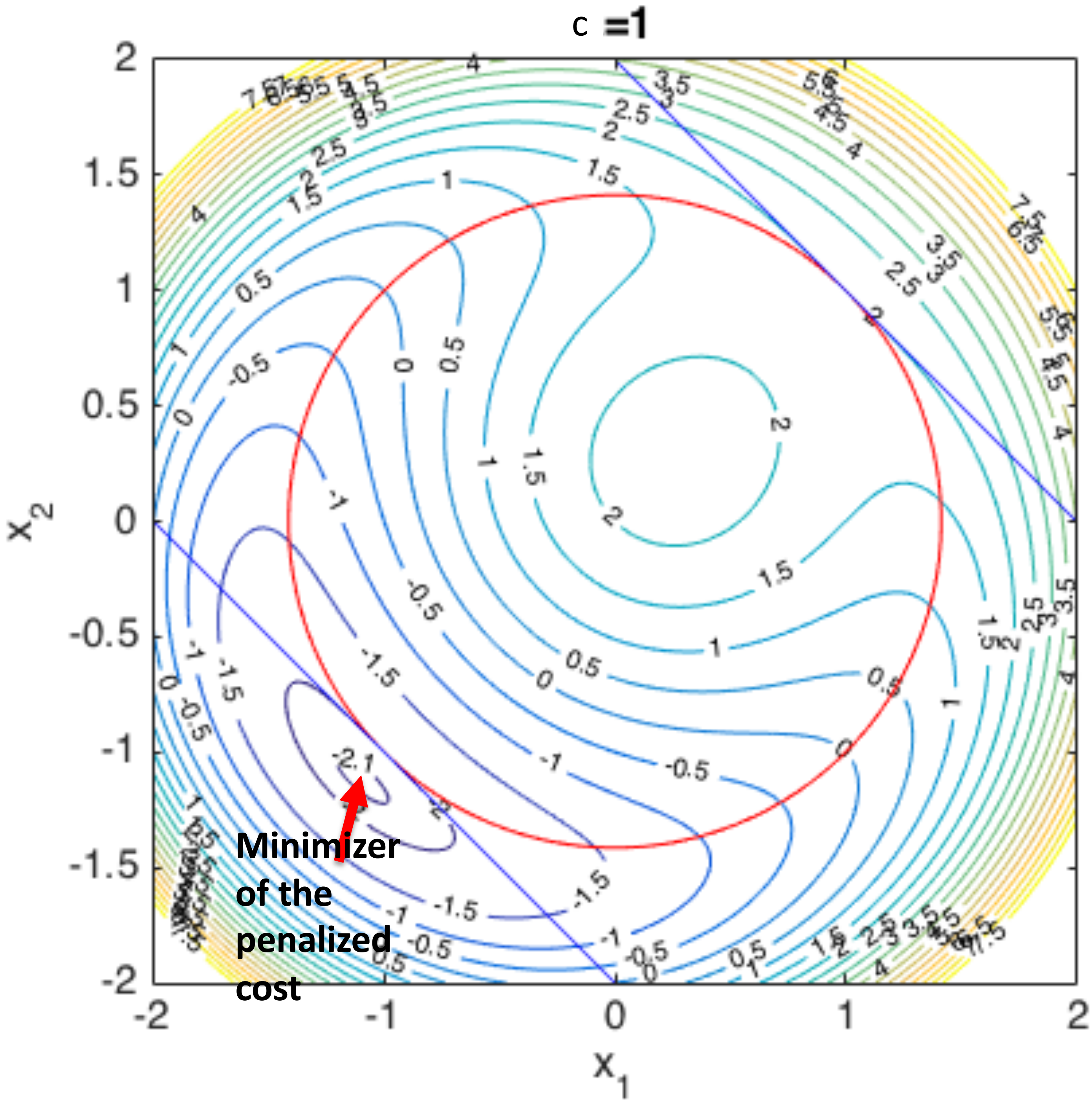
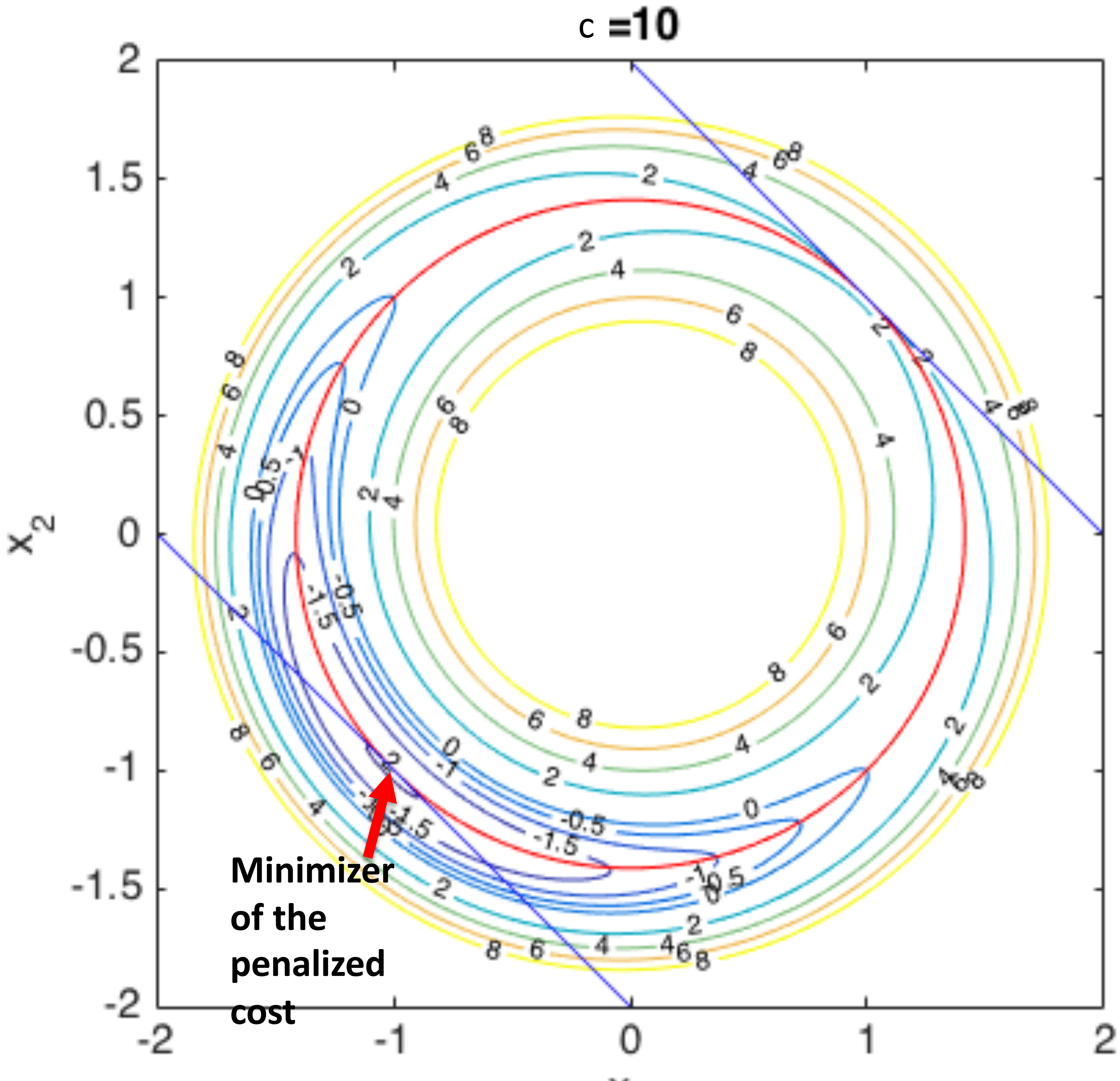


Figure 3



Penalty Method

The procedure for solving problem (1) by the penalty function method:

- Let $\{c_k\}$, $k = 1, 2, \ldots,$ be a sequence tending to infinity such that for each k, $c_k > 0$, and $c_{k+1} > c_k$.
- Define the function

$$Q(x; c)=f(x)+c\, P(x)$$

- For each k solve the problem

$$\text{Minimize } Q(x; c_k)$$

obtaining a solution c_k.

Convergence Guarantees of the Quadratic Penalty Method

Let $\bar{x}$ be the global minimizer of

minimize $f(x)$ subject to (3)

$h_i(x) = 0, \quad i=1,\ldots,m$

Suppose that each x_k is the exact global minimizer of $Q(x;\, c_k) = f(x) + \frac{c_k}{2}\sum_{i=1}^{m} h_i(x)^2$

for positive and monotonically increasing sequence of $\{c_k\}$ *where* $c_k \uparrow \infty$.

Then every limit point x^* of the sequence $\{x_k\}$ is a global solution of the constrained optimization problem (3).

Theorem

consider

$\min f(x)$ s.t. $h_i(x) = 0 \quad i = 1, \ldots, m$ $\quad$ ⊛ Let x^* be the global minimizer

consider the unconstrained penalized form

$$Q(x, c_k) = f(x) + \frac{c_k}{2} \sum_{i=1}^{m} h_i(x)^2 \qquad c_k > 0$$

Suppose that each x_k is the <u>exact global minimizer</u> of $Q(x, c_k)$, and that $c_k \to \infty$. Then every limit point $\bar{x}$ of sequence $\{x_k\}$ is a global solution of the problem ⊛

---

proof $\qquad x_k = \operatorname{argmin} Q(x, c_k) \Rightarrow \{x_k\} \xrightarrow[k \to \infty]{} \bar{x} = x^*$

x^* is a global minimizer of $(*)$

$$f(x^*) \leq f(x) \qquad x \in \Omega = \{x \in \mathbb{R}^n \mid h_i(x) = 0 \;\; i = 1, \ldots, m\}$$

x_k is a global minimizer of $Q(x, c_k)$

$$Q(x_k, c_k) \leq Q(x^*, c_k)$$

$$(**) \qquad f(x_k) + \frac{c_k}{2} \sum_{i=1}^{m} h_i(x_k)^2 \leq f(x^*) + \frac{c_k}{2} \sum_{i=1}^{m} \cancelto{0}{h_i(x^*)^2}$$

$$\sum_{i=1}^{m} h_i(x_k)^2 \leq \frac{2}{c_k} \left(f(x^*) - f(x_k) \right)$$

$$c_k \to \infty \Rightarrow \sum_{i=1}^{m} h_i(x_k)^2 \leq 0 \Rightarrow h_i(\bar{x}) = 0 \quad i = 1, \ldots, m$$

$$\{x_k\} \to \bar{x}$$

$c_k \to \infty$ $\quad$ $(**)$ $\Downarrow$

$$f(\bar{x}) + \frac{c_k}{2} \sum \cancelto{0}{h_i(\bar{x})^2} \leqslant f(x^*)$$

$$\begin{cases} f(\bar{x}) \leqslant f(x^*) \\ h_i(\bar{x}) = 0 \quad i = 1, \dots, m \\ \bar{x} \in \Omega \end{cases}$$

$\Rightarrow$ $\bar{x}$ is a global minimizer of the constraint optimization problem

ALGORITHMIC FRAMEWORK

A general framework for algorithms based on the quadratic penalty function can be specified as follows.

(Quadratic Penalty Method).

Given $c_0 > 0$, a nonnegative sequence $\{\tau_k\}$ with $\tau_k \to 0$, and a starting point x_0^s;
for $k = 0, 1, 2, \ldots$
  Find an approximate minimizer x_k of $Q(\cdot;\, c_k)$, starting at x_k^s,
    and terminating when $\|\nabla_x Q(x;\, c_k)\| \le \tau_k$;
  if final convergence test satisfied
    stop with approximate solution x_k;
  end (if)
  Choose new penalty parameter $c_{k+1} > c_k$;
  Choose new starting point x_{k+1}^s;
end (for)

The starting point x^s_{k+1} usually is selected to be x_k

Convergence Guarantees of the Practical Quadratic Penalty Method

Theorem- Suppose that the tolerances $\{\tau_k\}$ and penalty parameters $\{c_k\}$ satisfy $\tau_k \to 0$ and $c_k \uparrow \infty$. Then if a limit point x^* of the sequence $\{x_k\}$ is infeasible, it is a stationary point of the function $\|h(x)\|^2$. On the other hand, if a limit point x^* is feasible and the constraint gradients $\nabla h_i(x)$ are linearly independent, then x^* is a KKT point for the problem

$$\begin{cases} \text{minimize } f(x) & \text{subject to} \\ \mathrm{h_i}(x) = 0, & \text{i=1,...,m} \end{cases}$$

For such points, we have for any infinite subsequence K such that $\lim_{k\in K} x_k = x^*$ that

$$\lim_{k\in K} c_k h_i(x_k) = \lambda_i^* \qquad i = 1,\ldots,m$$

where λ_i^* is the multiplier vector that satisfies the KKT conditions (first order necessery conditions for optimality) for the equality constrained problem.

Exact Penalty functions

It is possible to construct penalty functions that are exact in the sense that

- the solution of the penalty problem yields the exact solution to the original problem for a finite value of the penalty parameter.

- With these functions <u>it is not</u> necessary to solve an infinite sequence of penalty problems to obtain the correct solution.

- Difficulty: these penalty functions are non-differentiable.

Exact Penalty functions

$$\boxed{\begin{array}{ll} \text{minimize } f(x) & \text{subject to} \\ \mathrm{h_i}(x)=0, & \text{i=1,...,m} \\ g_{\mathrm{j}}(x)\le 0, & \text{j=1,...,r} \end{array}}$$

$$\Downarrow$$

$$\text{minimize } f(x) + \underbrace{\mu \sum_{i=1}^{m} \left|\mathrm{h_i}(x)\right| + \text{ c } \sum_{j=1}^{r} \max\{0, g_j(x)\}}_{\text{c P(x)}}$$

Here, the solution of the penalty problem yields the exact solution to the original problem for a finite value of the penalty parameter c.

Exact penalty function theorem: Suppose the point x^* satisfies the second-order sufficiency conditions for a local minimum of the constrai[nt] problem. Let λ and μ be corresponding Lagrange multiplier vectors then for

$$c \geq \max\left(\max\{|\lambda_i|\}_{i=1}^{m},\ \max\{\mu_j\}_{j=1}^{p} \right)$$

x^* is also a local minimum of the absolute-value penalty objective $\min f(x) + cP(x)$

Example $\min\ 2x^2 + 2xy + y^2 - 2y$ / $\min\ y^2 - 2y$

s.t. $x = 0$

$x^* = 0,\ y^* = 1,\ \lambda^* = -2$

* $Q(x,c) = 2x^2 + 2xy + y^2 - 2y + cx^2$

$\left(x_c^* = -\frac{2}{2+c},\ y_c^* = 1 - \frac{2}{2+c}\right) \xRightarrow{c \to \infty}$

$\left(x_c^* \to 0,\ y_c^* \to 1\right)$

* $Q_e(x,c) = 2x^2 + 2xy + y^2 - 2y + c|x|$

$$\underbrace{x^2 + \underbrace{(2x + c|x|)}_{c \geq 2}}_{x^* = 0} + \underbrace{(y - 1 + x)^2}_{y^* = 1} - 1$$

for any $c \geq 2$ the ~~optim~~ minimizer of $Q_e(x,c)$ and the origional problem are the same

$$c \geq \|\lambda^*\|_\infty = \max\{|\lambda_i|\}_{i=1}^m$$

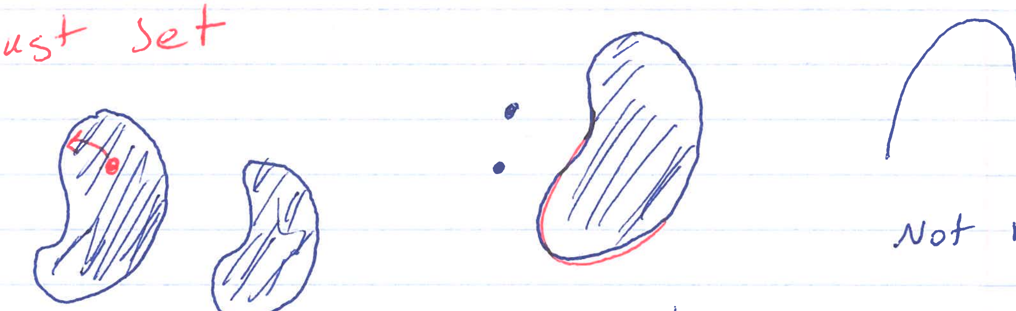
Barrier Methods

$$\begin{cases} \min & f(x) \\ \text{subject to} & x \in \Omega \end{cases}$$

Ω: has a non-empty interior point.

↳ the set has an interior and it is possible to get to any boundary point by approaching it from the interior.

Robust set



Robust

Not robust

Not robust

$$f(x) + c\,B(x)$$

Barrier

Example: Let $\Omega = \{ x \mid g_i(x) \leq 0 \quad i = 1, \dots, r \}$

$$B(x) = -\sum_{i=1}^{r} \frac{1}{g_i(x)}$$

Logarithmic utility function

$$B(x) = -\sum_{i=1}^{r} \log[-g_i(x)]$$