

Lecture 10 - Penalty and Augmented Lagrangian methods

Dinesh Krishnamoorthy

Department of Engineering Cybernetics, NTNU

Spring 2026

Outline

- ▶ Constrained Optimization problem
- ▶ Transform constrained optimization problem into a sequential unconstrained minimization problem
- ▶ Lagrangian Duality and Dual Ascent
- ▶ Augmented Lagrangian methods

Sources

Presentation based on material from:

- ▶ Nocedal & Wright (2006): Numerical optimization (Chapters 15, 16, 17, 18)

Objectives and Readings

Learning Objectives:

1. *understand* interior penalty, exterior penalty, and augmented Lagrangian transformation methods for sequential unconstrained minimization.
2. *understand* the principles of dual ascent, Lagrangian decomposition, and alternating directions method of multipliers.
3. *understand* how to decompose a large-scale optimization problem.
4. *can demonstrate* the use of such algorithms for distributed learning problems

Readings: Sects. 5.5, 6.4, 6.5, 6.6, and 6.7, Nocedal & Wright (2006): Numerical optimization

Constrained problem

$$\begin{array}{ll}\min_{x \in \mathbb{R}^n} & f(x) \\ \text{subject to} & g(x) \leq 0 \\ & h(x) = 0\end{array}$$

- ▶ Smoothness: $f(x)$, $g(x)$, $h(x)$ twice continuously differentiable
- ▶ $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_g}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_h}$

Handling constraint

Constrained Optimization

$$\min_x f(x)$$

$$\text{s. t. } g(x) \leq 0$$

$$h(x) = 0$$

transform into
unconstrained
problem



Penalty methods

$$\min_x T(x, r) := f(x) + p(g(x), h(x), r)$$



Direct Methods

$$\underbrace{\begin{bmatrix} \nabla f(x^*) + \lambda^{*T} \nabla h(x^*) + \mu_{\mathbb{A}}^{*T} \nabla g_{\mathbb{A}}(x^*) \\ h(x^*) \\ g_{\mathbb{A}}(x^*) \end{bmatrix}}_{:= F(x, \lambda, \mu)} = 0$$

Transformation: penalty function methods

- Transform constrained problem into unconstrained one:

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) \leq 0 \end{array} \quad \rightarrow \quad \min_{x \in \mathbb{R}^n} T(x, r) = f(x) + p(g(x), h(x), r)$$

Transformation: penalty function methods

- Transform constrained problem into unconstrained one:

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) \leq 0 \end{array} \quad \rightarrow \quad \min_{x \in \mathbb{R}^n} T(x, r) = f(x) + p(g(x), h(x), r)$$

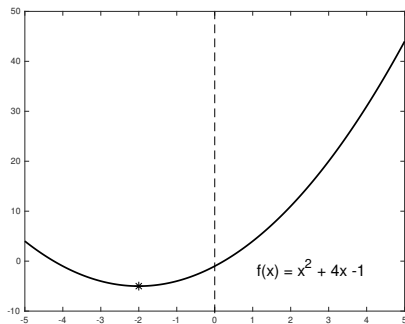
- Penalty function p :
 - Exterior penalty functions / Quadratic penalty functions
 - Interior penalty functions (or barrier functions)
 - Augmented Lagrangian penalty functions

Exterior penalty function

Equality constrained problem

$$\min_x f(x) \quad \text{s.t. } h(x) = 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|h(x)\|^2$$

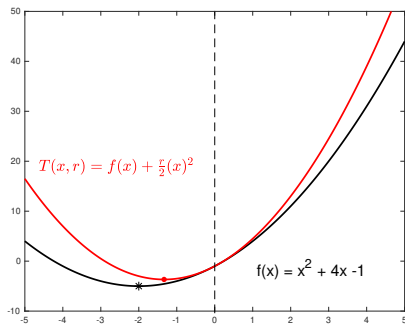


Exterior penalty function

Equality constrained problem

$$\min_x f(x) \quad \text{s.t. } h(x) = 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|h(x)\|^2$$

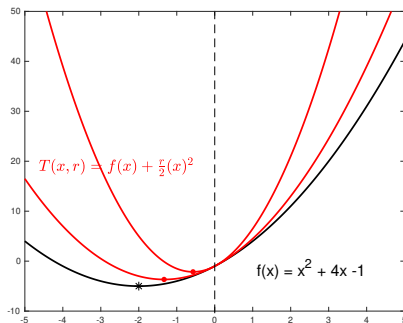


Exterior penalty function

Equality constrained problem

$$\min_x f(x) \quad \text{s.t. } h(x) = 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|h(x)\|^2$$

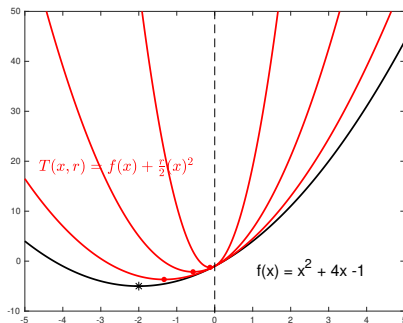


Exterior penalty function

Equality constrained problem

$$\min_x f(x) \quad \text{s.t. } h(x) = 0$$

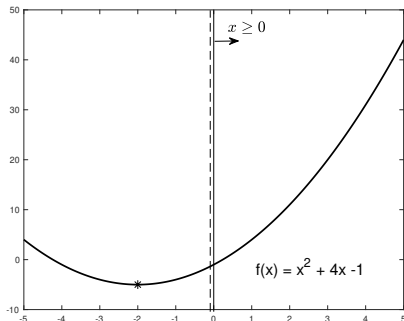
$$\min_x T(x, r) = f(x) + \frac{r}{2} \|h(x)\|^2$$



Exterior penalty function

$$\min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|[g(x)]^-\|^2$$

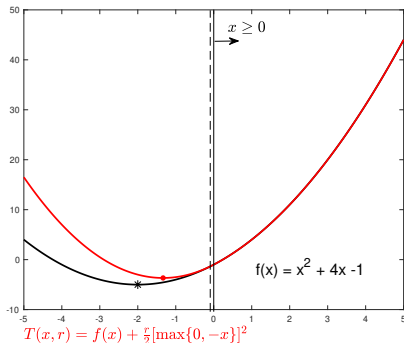


where the notation $[g(x)]^- = \max\{0, g(x)\}$

Exterior penalty function

$$\min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|[g(x)]^-\|^2$$

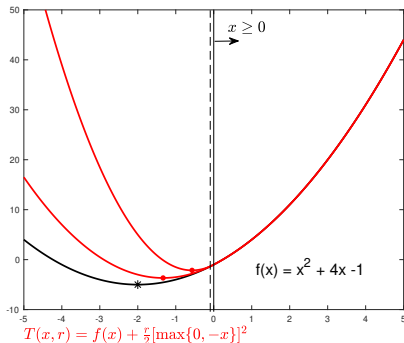


where the notation $[g(x)]^- = \max\{0, -g(x)\}$

Exterior penalty function

$$\min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|[g(x)]^-\|^2$$

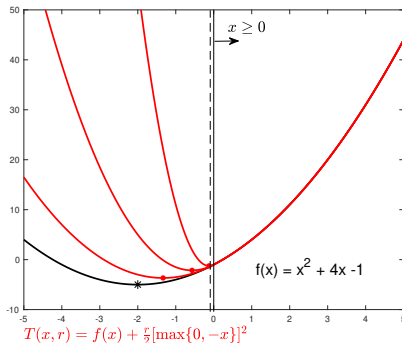


where the notation $[g(x)]^- = \max\{0, g(x)\}$

Exterior penalty function

$$\min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|[g(x)]^-\|^2$$

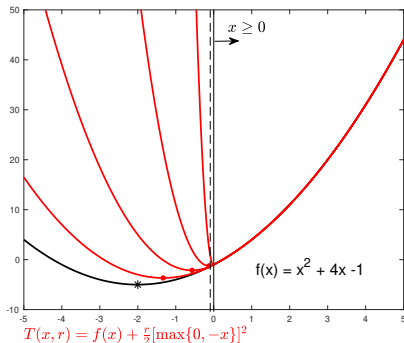


where the notation $[g(x)]^- = \max\{0, -g(x)\}$

Exterior penalty function

$$\min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\min_x T(x, r) = f(x) + \frac{r}{2} \|[g(x)]^-\|^2$$



where the notation $[g(x)]^- = \max\{0, g(x)\}$

Exterior penalty function

- Constrained problem:

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & h_j(x) = 0 \quad j = 1, \dots, m_h \\ & g_j(x) \leq 0 \quad j = 1, \dots, m_g \end{array}$$

- Unconstrained exterior penalty transformation:

$$\min_{x \in \mathbb{R}^n} T(x, r) = f(x) + \frac{r}{2} \sum_{j=1}^{m_h} (h_j(x))^2 + \frac{r}{2} \sum_{j=1}^{m_g} ([g_j(x)]^-)^2$$

- Solve for increasing sequence $\{r_k\}$ such that:

$$r_1 < r_2 < \dots < r_k \quad , \quad \lim_{k \rightarrow \infty} r_k \rightarrow \infty$$

Intuition: Penalize constraint violation with increased severity

Exercise

Consider optimization problem:

$$\begin{array}{ll}\min_x & f(x) = \frac{1}{4}x^2 \\ \text{s.t.} & g(x) = 2 - x \leq 0 \\ & \chi : x > 0\end{array}$$

Calculate analytically the minimizer of the transformed function using quadratic penalty for $r = 0.1$, $r = 2$, $r = 10$, and r in the limit approaching ∞ .

Interior penalty function

What happens if your inequality constraint is a **safety-critical constraint**?

- Constrained problem (inequality only):

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & g_j(x) \leq 0 \quad j = 1, \dots, m_g \end{array}$$

Interior penalty function

What happens if your inequality constraint is a **safety-critical constraint**?

- Constrained problem (inequality only):

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & g_j(x) \leq 0 \quad j = 1, \dots, m_g \end{array}$$

Intuition: Cautiously approach the constraint by penalizing getting too close to the constraint.

Interior penalty function

- Constrained problem (inequality only):

$$\begin{array}{ll}\min_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_j(\mathbf{x}) \leq 0 \quad j = 1, \dots, m_g\end{array}$$

- Unconstrained interior penalty transformation:

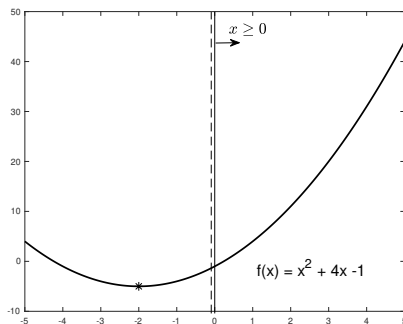
inverse:
$$\min_{\mathbf{x} \in \mathbb{R}^n} T(\mathbf{x}, \tau) = f(\mathbf{x}) - \tau \sum_{j=1}^{m_g} \frac{1}{g_j(\mathbf{x})}$$

logarithmic:
$$\min_{\mathbf{x} \in \mathbb{R}^n} T(\mathbf{x}, \tau) = f(\mathbf{x}) - \tau \sum_{j=1}^{m_g} \ln [-g_j(\mathbf{x})]$$

- Solve for decreasing sequence $\{\tau_k\}$
- Log-barrier $\rightarrow$ Interior-point methods for LP, QP, and NLP to treat the inequality constraints

Inverse Interior penalty function

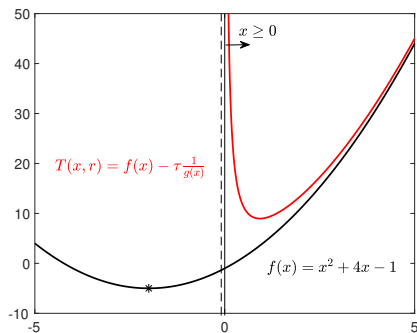
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \frac{1}{g_j(x)}$$



$$\tau = 0$$

Inverse Interior penalty function

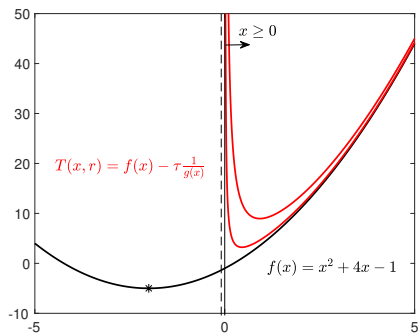
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \frac{1}{g_j(x)}$$



$$\tau = 5$$

Inverse Interior penalty function

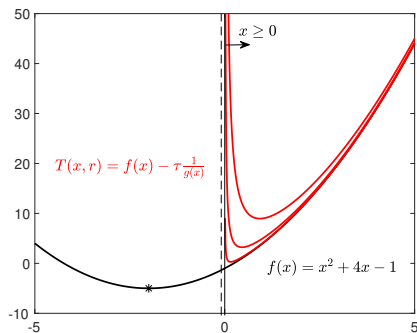
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \frac{1}{g_j(x)}$$



$$\tau = 1$$

Inverse Interior penalty function

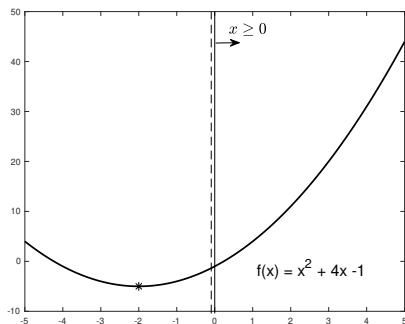
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \frac{1}{g_j(x)}$$



$$\tau = 0.1$$

Barrier penalty function

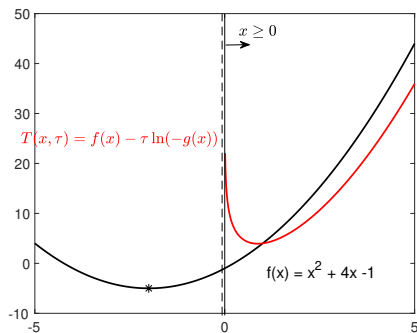
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \ln [-g_j(x)]$$



$$\tau = 0$$

Barrier penalty function

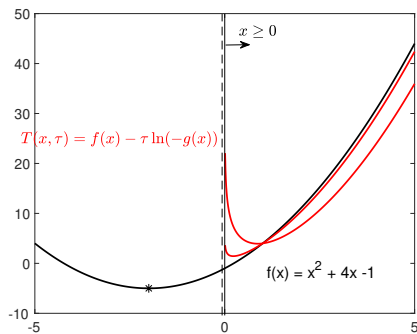
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \ln [-g_j(x)]$$



$$\tau = 5$$

Barrier penalty function

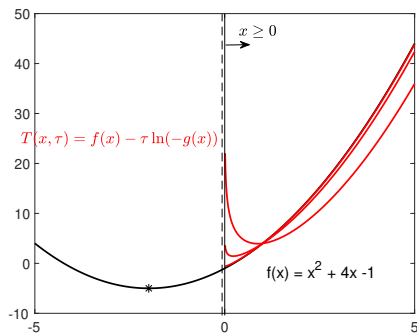
$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \ln [-g_j(x)]$$



$$\tau = 1$$

Barrier penalty function

$$\min_{x \in \mathbb{R}^n} T(x, \tau) = f(x) - \tau \sum_{j=1}^{m_g} \ln [-g_j(x)]$$



$$\tau = 0.1$$

Exercise

Consider the same optimization problem:

$$\begin{array}{ll}\min_x & f(x) = \frac{1}{4}x^2 \\ \text{s.t.} & g(x) = 2 - x \leq 0 \\ & \chi : x > 0\end{array}$$

Calculate analytically the minimizer of the transformed function using log-barrier function for $\tau = 1$, $\tau = 0.5$, $\tau = 0.1$, and τ in the limit approaching zero.

Initial point when using iterative methods

Iterative methods for unconstrained optimization problem $\min_x T(x, \tau)$
use

$$x_{k+1} = x_k + \alpha_k s_k$$

starting from the initial guess x_0 .

Initial point when using iterative methods

Iterative methods for unconstrained optimization problem $\min_x T(x, \tau)$
use

$$x_{k+1} = x_k + \alpha_k s_k$$

starting from the initial guess x_0 .

- ▶ **Question:** What x_0 can you use when using exterior penalty methods?
- ▶ **Question:** What x_0 can you use when using interior penalty methods?

Notes on penalty functions

- ▶ Ill-conditioning for large r values (exterior penalty) and small τ values (interior penalty)
- ▶ Interior and exterior penalty functions are inexact (unconstrained penalty minimizer $\neq$ constrained minimizer)
- ▶ Well-known exact penalty method is l_1 -penalty:

$$\min_{x \in \mathbb{R}^n} T(x, r) = f(x) + r \sum_{j=1}^{m_h} |h_j(x)| + r \sum_{j=1}^{m_g} [g_j(x)]^-$$

which is non-smooth, but has minimizer equal to constrained problem for sufficiently large r

- ▶ Merit function:
 ℓ_1 penalty function is commonly employed as measure of improvement in line search and trust region methods for constrained optimization

Augmented Lagrangian (multiplier) methods

Motivation:

- ▶ Avoid ill-conditioning without introducing non-smoothness

Conceptual idea:

- ▶ Add quadratic penalty to Lagrangian function $\mathcal{L}(x, \lambda, \mu)$
- ▶ Update estimates of multipliers λ_k, μ_k
- ▶ Successively decrease r_k for efficiency
(convergence occurs for any positive value of r_k)

Duality

$$V^* = \min_x f(x) \quad \text{s.t. } h(x) = 0$$

The Lagrangian of this problem is

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T h(x)$$

Define the dual objective function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

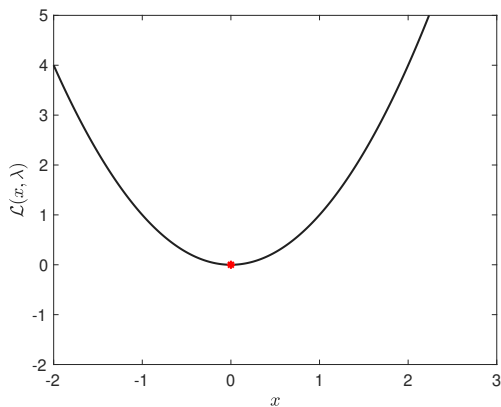
Dual problem

$$\begin{aligned} & \max_{\lambda} q(\lambda) \\ \Rightarrow V^* &= \max_{\lambda} \min_x \mathcal{L}(x, \lambda) \end{aligned}$$

Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

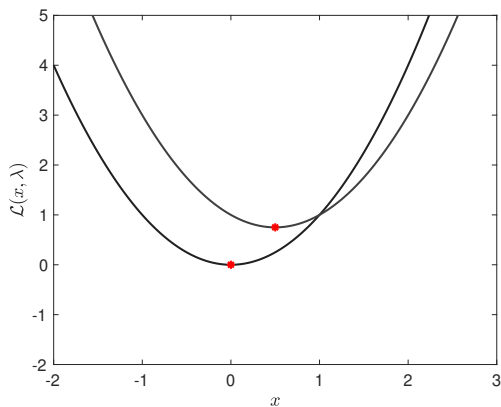
hold $\lambda = 0$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

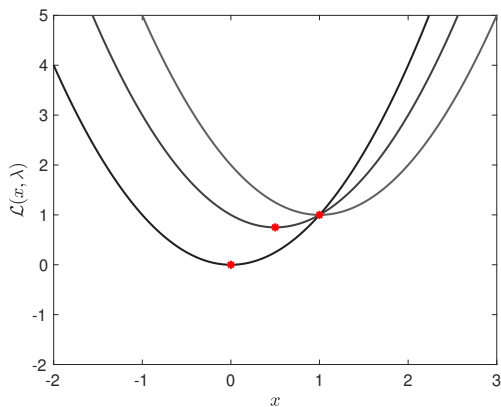
hold $\lambda = 1$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

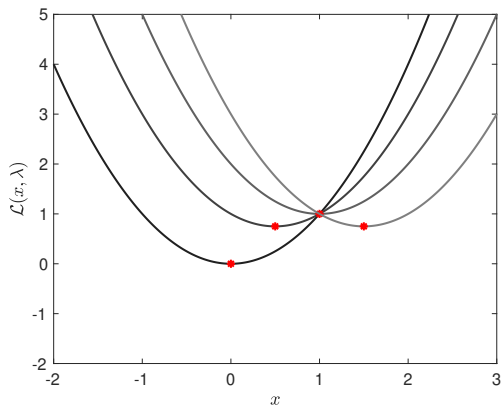
hold $\lambda = 2$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

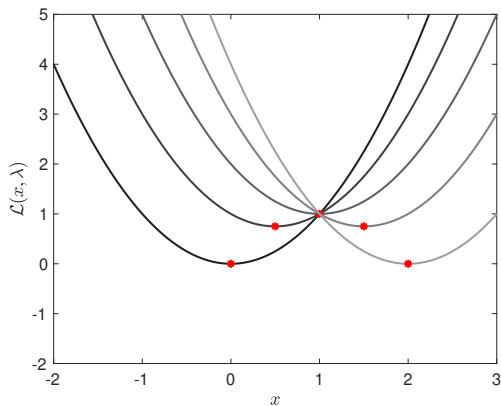
hold $\lambda = 3$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

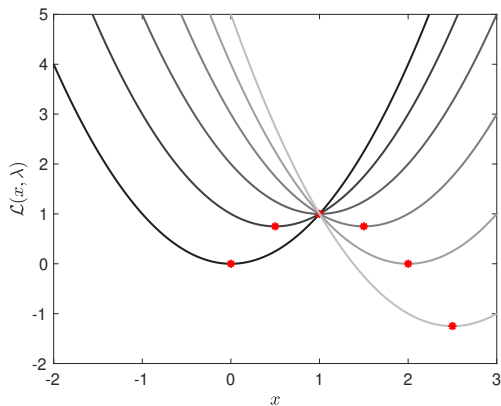
hold $\lambda = 4$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

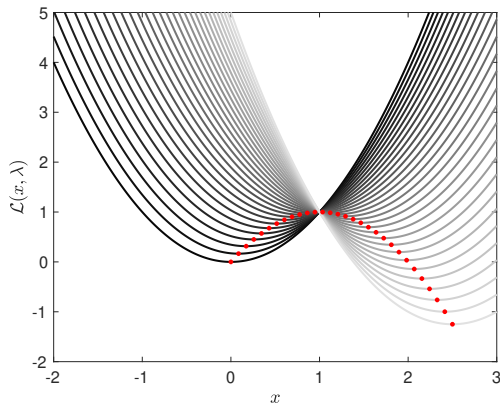
hold $\lambda = 5$



Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

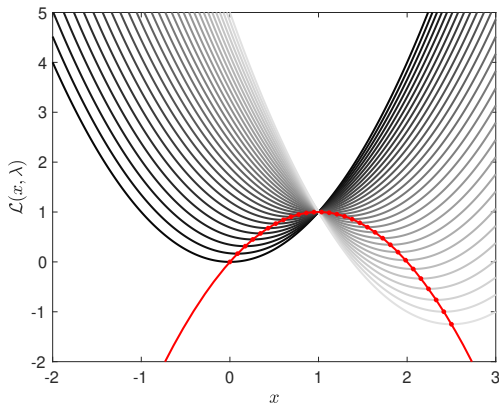
hold λ between $[0, 5]$



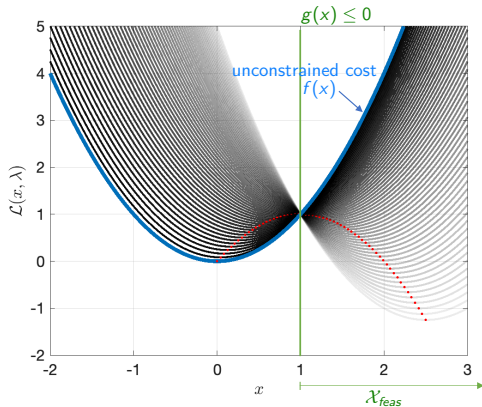
Duality

$$\mathcal{L}(x, \lambda) = x^2 + \lambda(1 - x)$$

$$q(\lambda) = \inf_x \mathcal{L}(x, \lambda)$$



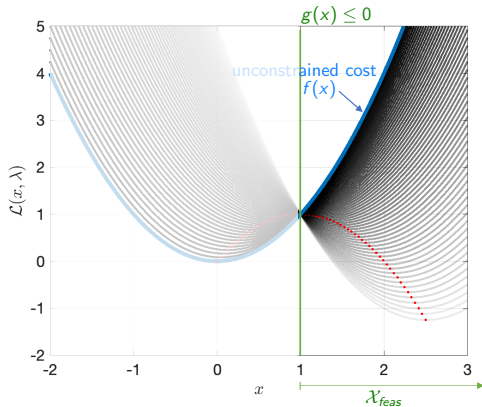
Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

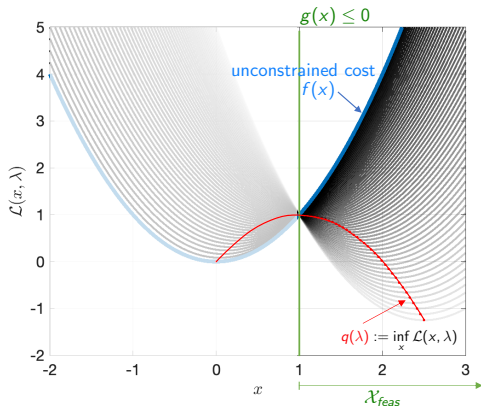
Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Duality



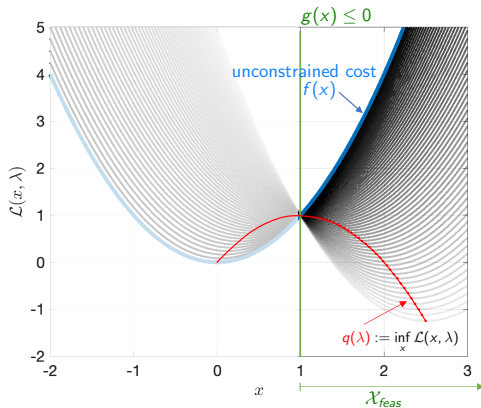
$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

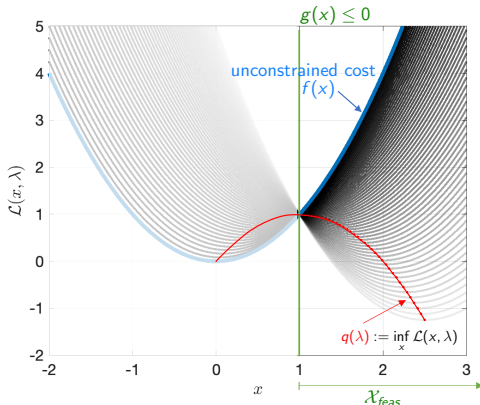
Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

$q(\lambda)$ is always concave and has the property

$$V^* \geq q(\lambda)$$

Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

$q(\lambda)$ is always concave and has the property

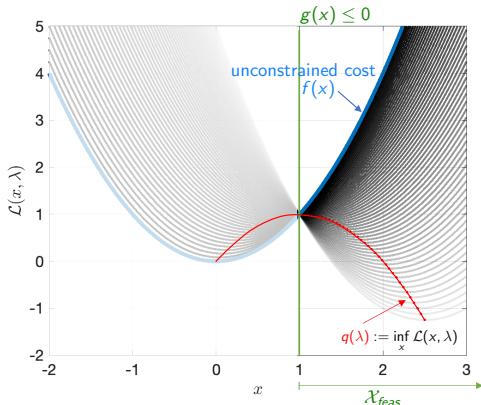
$$V^* \geq q(\lambda)$$

Proof: For any feasible $x \in \mathcal{X}_{feas}$ and $\lambda \geq 0$

$$f(x) \geq \mathcal{L}(x, \lambda)$$

Why? Because for any feasible x , $g(x) \leq 0$ (-ve) and $\lambda \geq 0$ (+ve)

Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

$q(\lambda)$ is always concave and has the property

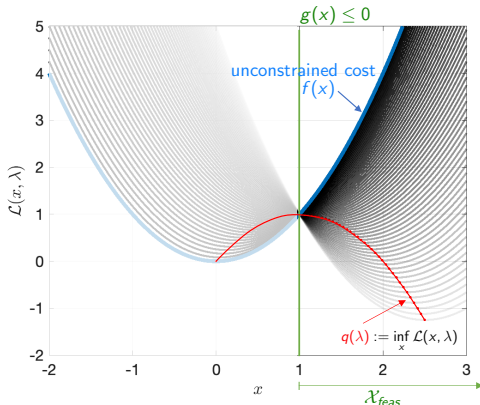
$$V^* \geq q(\lambda)$$

Proof: For any feasible $x \in \mathcal{X}_{feas}$ and $\lambda \geq 0$

$$\Rightarrow f(x) \geq \mathcal{L}(x, \lambda) \geq \inf_{x \in \mathcal{X}_{feas}} \mathcal{L}(x, \lambda) = q(\lambda)$$

At the optimal $x^* \in \mathcal{X}_{feas}$, $f(x^*) = V^*$, which proves our lower bound property

Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

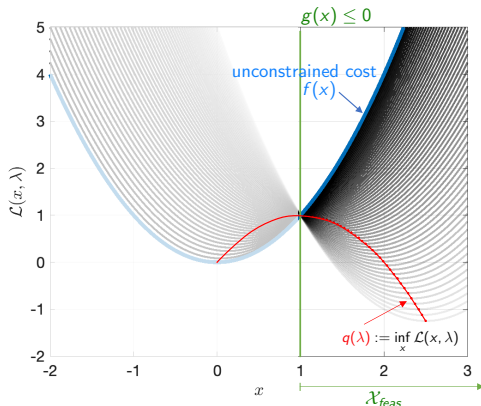
$q(\lambda)$ is always concave and has the property

$$V^* \geq q(\lambda)$$

Now that we have proven, that $q(\lambda)$ is always lower than the optimal cost V^* , in order to get the best lower bound (i.e. get as close to V^* as possible), we need to maximize the dual function.

$$\Rightarrow \max_{\lambda} q(\lambda)$$

Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

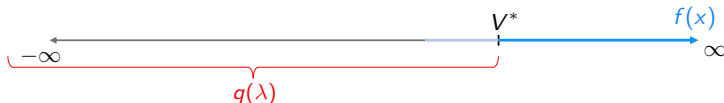
$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

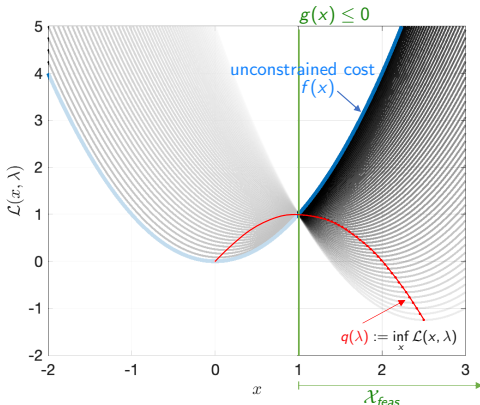
$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

$q(\lambda)$ is always concave and has the property

$$V^* \geq q(\lambda)$$



Duality



$$V^* = \min_x f(x) \quad \text{s.t. } g(x) \leq 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$$

Define the dual function

$$q(\lambda) := \inf_x \mathcal{L}(x, \lambda)$$

$q(\lambda)$ is always concave and has the property

$$V^* \geq q(\lambda)$$

Why infimum in $q(\lambda) = \inf_x \mathcal{L}(x, \lambda)$?

For some very large λ in the feasible region, $\mathcal{L}(x, \lambda) \rightarrow -\infty$. Infimum of $\mathcal{L}(x, \lambda)$ denotes to a point which is smaller than every other value $\mathcal{L}(x, \lambda)$ can take for some very large $\lambda \rightarrow \infty$.

Exercise

Exercise 4

Consider the problem

$$\min_{x_1, x_2} \frac{1}{2}(x_1^2 + x_2^2) \quad \text{s.t.} \quad 1 - x_1 \leq 0$$

- a) Write the dual problem $q(\lambda)$ for this optimization problem.
- b) Compute the optimal Lagrange multiplier by maximizing the dual problem.

Dual Ascent

$$\min_x f(x) \quad \text{s.t.} \quad h(x) = 0$$

$$\max_{\lambda} \min_x \mathcal{L}(x, \lambda) = f(x) + \lambda^T h(x)$$

Intuition behind dual ascent.

- ▶ Two variables: x and λ .
- ▶ Fix one and solve for the other

Dual Ascent

$$\min_x f(x) \quad \text{s.t.} \quad h(x) = 0$$

$$\max_{\lambda} \min_x \mathcal{L}(x, \lambda) = f(x) + \lambda^T h(x)$$

Dual Ascent

Starting from some initial guess for λ_0

For $k = 0, 1, 2, \dots$

- ▶ $x_{k+1} = \arg \min_x \mathcal{L}(x, \lambda_k)$
- ▶ $\lambda_{k+1} = \arg \max_{\lambda} \mathcal{L}(x_{k+1}, \lambda)$

Dual Ascent

$$\min_x f(x) \quad \text{s.t.} \quad h(x) = 0$$

$$\max_{\lambda} \min_x \mathcal{L}(x, \lambda) = f(x) + \lambda^T h(x)$$

Dual Ascent

Starting from some initial guess for λ_0

For $k = 0, 1, 2, \dots$

- ▶ $x_{k+1} = \arg \min_x f(x) + \lambda_k^T h(x)$
- ▶ $\lambda_{k+1} = \lambda_k + \alpha_k h(x_{k+1})$

Augmented Lagrangian (multiplier) methods

- Constrained problem (equality constraints):

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad h(x) = 0$$

Augmented Lagrangian (multiplier) methods

- Constrained problem (equality constraints):

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad h(x) = 0$$

- Augmented Lagrangian function

$$x_{k+1} = \arg \min \mathcal{L}(x, \lambda_k, \rho_k) = f(x) + \lambda_k^\top h(x) + \frac{\rho_k}{2} \|h(x)\|^2$$

Augmented Lagrangian (multiplier) methods

- Constrained problem (equality constraints):

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{s.t.} \quad h(\mathbf{x}) = 0$$

- Augmented Lagrangian function

$$\mathbf{x}_{k+1} = \arg \min \mathcal{L}(\mathbf{x}, \lambda_k, \rho_k) = f(\mathbf{x}) + \lambda_k^\top h(\mathbf{x}) + \frac{\rho_k}{2} \|h(\mathbf{x})\|^2$$

- Lagrange multiplier update (1st order):

$$\lambda_{k+1} = \lambda_k + \rho_k h(\mathbf{x}_{k+1})$$

Augmented Lagrangian (multiplier) methods

- Constrained problem (equality constraints):

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{s.t.} \quad h(\mathbf{x}) = 0$$

- Augmented Lagrangian function

$$\mathbf{x}_{k+1} = \arg \min \mathcal{L}(\mathbf{x}, \lambda_k, \rho_k) = f(\mathbf{x}) + \lambda_k^\top h(\mathbf{x}) + \frac{\rho_k}{2} \|h(\mathbf{x})\|^2$$

- Lagrange multiplier update (1st order):

$$\lambda_{k+1} = \lambda_k + \rho_k h(\mathbf{x}_{k+1})$$

- Penalty parameter update: $\rho_{k+1} \geq \rho_k$.

Notes on Augmented Lagrangian (multiplier) methods

- ▶ Inequality constraints may be converted into equality constraints using slack variables: $g_j + z_j = 0$ with $z_j \geq 0$ to arrive at a bound constrained transformed problem. The bounds may be treated efficiently using a projection method.
(see Nocedal and Wright, 2006, Sect 17.4)
- ▶ Augmented Lagrangian penalty function is also employed as merit function in SQP (as alternative to l_1 penalty).
- ▶ Augmented Lagrangian multiplier methods are used in distributed optimization to coordinate the coupling between decomposed optimization subproblems.

Example

Consider optimization problem:

$$\begin{array}{ll} \min_x & f(x) = \frac{1}{4}x^2 \\ \text{s.t.} & g(x) = 2 - x \leq 0 \end{array}$$

- a) Give an analytical expression for the minimizer x^* of the transformed function in terms of μ and ρ
- b) Start with $\rho = 1$ and Lagrange multiplier estimate $\mu = 0$, and calculate the minimizer.
- c) Update μ ; update ρ using the factor $\rho^{k+1} = 10\rho^k$; and subsequently re-calculate the minimizer.

Summary

Constrained Optimization

$$\min_x f(x)$$

$$\text{s. t. } g(x) \leq 0$$

$$h(x) = 0$$

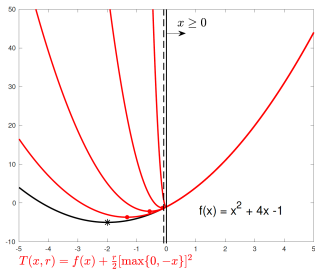
transform into
unconstrained
problem

Penalty methods

$$\min_x T(x, r) := f(x) + p(g(x), h(x), r)$$

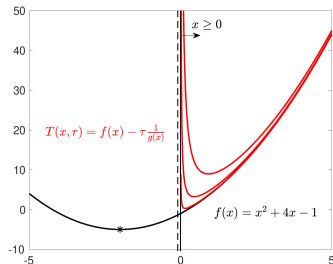
Exterior penalty

Sequentially increase severity



Interior penalty

Sequentially decrease caution



Summary

Constrained Optimization

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s. t. } & h(x) = 0 \end{aligned}$$

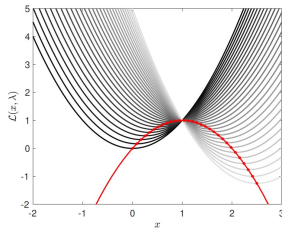
Lagrangian Duality

$$\begin{aligned} \max_{\lambda} \min_x \mathcal{L}(x, \lambda) &= f(x) + \lambda^T h(x) \\ \max_{\lambda} q(\lambda) &= \inf_x \mathcal{L}(x, \lambda) \end{aligned}$$

For $k = 0, 1, 2, \dots$

- $x^{k+1} = \arg \min_x f(x) + \lambda^{kT} h(x)$
- $\lambda^{k+1} = \lambda^k + \alpha^k h(x^{k+1})$

} Dual Ascent



Augmented Lagrangian

$$\mathcal{L}(x, \lambda, \rho) = f(x) + \lambda^T h(x) + \frac{\rho}{2} \|h(x)\|^2$$

For $k = 0, 1, 2, \dots$

- $x^{k+1} = \arg \min_x f(x) + \lambda^{kT} h(x) + \frac{\rho^k}{2} \|h(x)\|^2$
- $\lambda^{k+1} = \lambda^k + \rho^k h(x^{k+1})$