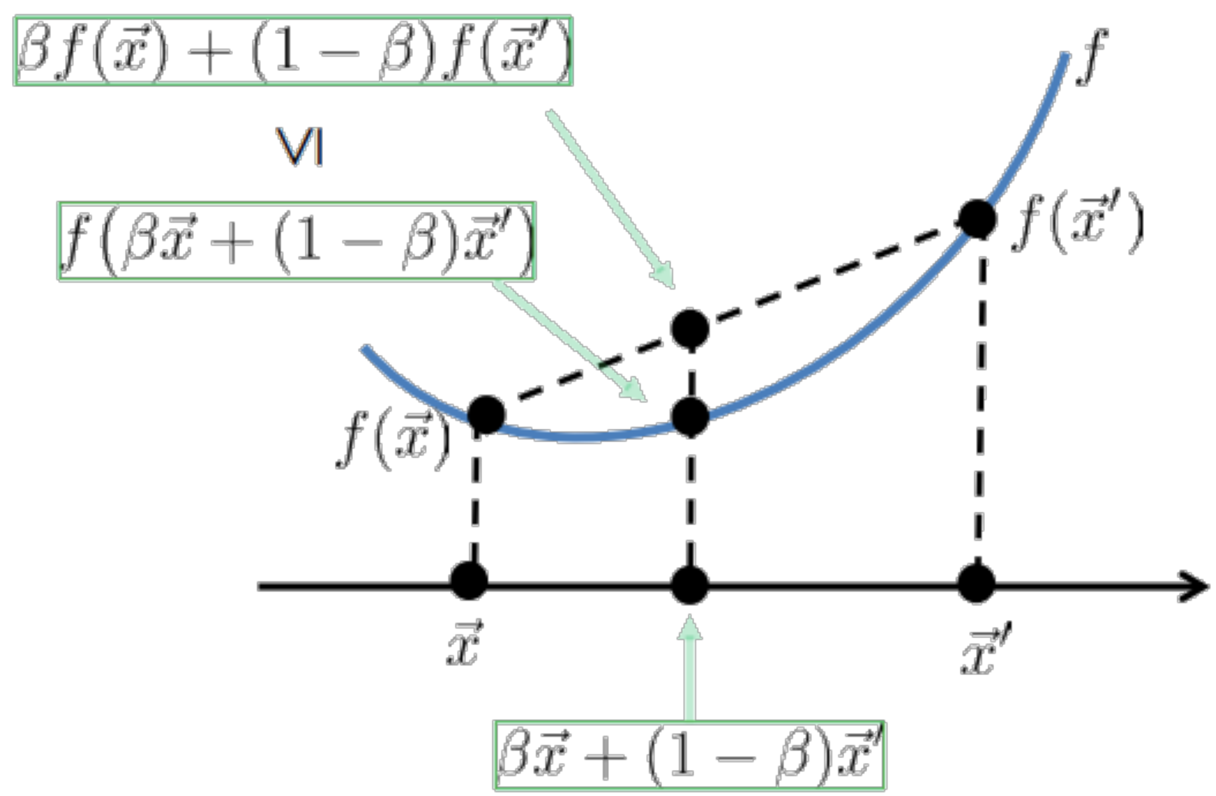
Constraint Optimization Primer

Convexity

A function f: $\mathbf{R}^d \rightarrow \mathbf{R}$ is called convex iff for any two points x, x' and $\beta \in [0,1]$

$$f\big(\beta\vec{x} + (1-\beta)\vec{x}'\big) \leq \beta f(\vec{x}) + (1-\beta) f(\vec{x}')$$

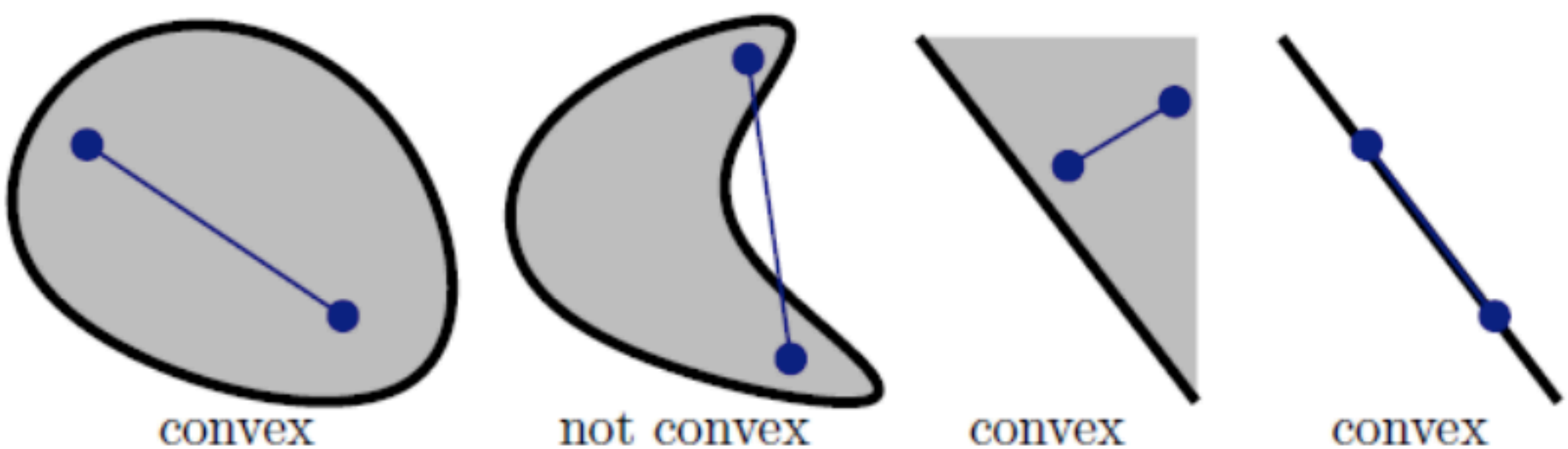


Convexity

A set $S \subset \mathbf{R}^d$ is called convex iff for any two points $x, x' \in S$ and any $\beta \in [0,1]$

$$\beta\vec{x} + (1-\beta)\vec{x}' \in S$$

Examples:



Convex Optimization

A constrained optimization

$$\underset{\vec{x} \in \mathbf{R}^d}{\text{minimize}} \quad f(\vec{x}) \qquad \textit{(objective)}$$

$$\text{subject to:} \quad g_i(\vec{x}) \leq 0 \quad \text{for } 1 \leq i \leq n \qquad \textit{(constraints)}$$

is called convex a convex optimization problem

If:

the objective function $f(\vec{x})$ is convex function, and

the feasible set induced by the constraints g_i is a convex set

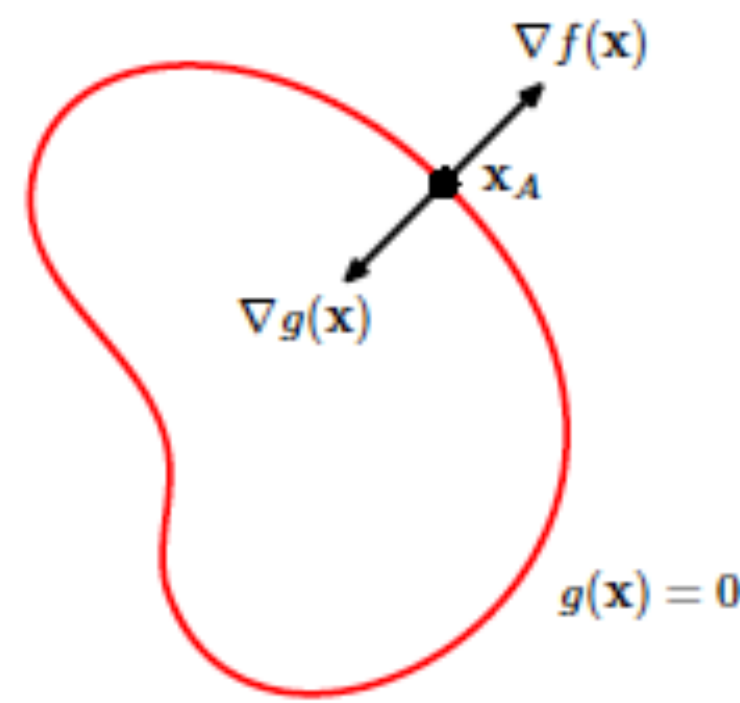
Why do we care?

*We and find the optimal solution for convex problems **efficiently**!*

Lagrangian

maximum of a function $f(x_1, x_2)$

$$g(x_1, x_2) = 0.$$



- ∇g is normal to the surface

$$g(\mathbf{x} + \epsilon) \simeq g(\mathbf{x}) + \epsilon^{\mathrm{T}} \nabla g(\mathbf{x}).$$

- ∇f(x) is also orthogonal to the constraint surface

- Lagrangian function $\nabla f + \lambda \nabla g = 0$

$$L(\mathbf{x}, \lambda) \equiv f(\mathbf{x}) + \lambda g(\mathbf{x}). \qquad \nabla_{\mathbf{x}} L = 0. \qquad \partial L / \partial \lambda = 0$$

-

Example

$$f(x_1, x_2) = 1 - x_1^2 - x_2^2$$

$$g(x_1, x_2) = x_1 + x_2 - 1 = 0,$$

$$L(\mathbf{x}, \lambda) = 1 - x_1^2 - x_2^2 + \lambda(x_1 + x_2 - 1).$$

$$\begin{aligned} -2x_1 + \lambda &= 0 \\ -2x_2 + \lambda &= 0 \\ x_1 + x_2 - 1 &= 0. \end{aligned}$$

$$(x_1^\star, x_2^\star) = (\tfrac{1}{2}, \tfrac{1}{2}).$$

Lagrangian

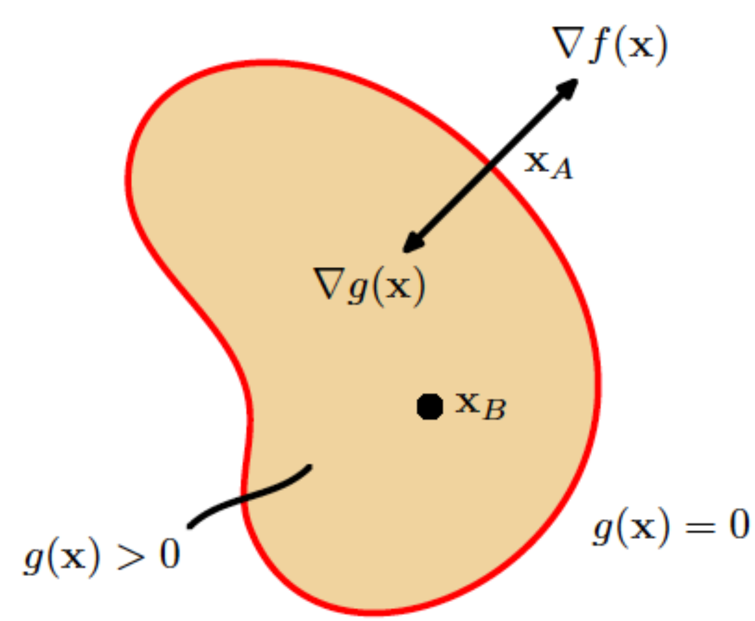
$f(\mathbf{x})$ subject to an *inequality constraint* of the form $g(\mathbf{x}) \geqslant 0$.

- Inactive : constrained stationary point lies in the region where g(x) > 0, g(x) plays no role, stationary condition is simply ∇f(x) = 0, stationary point of the Lagrange function with λ = 0.
- Active : it lies on the boundary g(x) = 0, corresponds to a stationary point of the Lagrange function with λ != 0.
- f(x) will only be at a maximum if its gradient is oriented away from the region g(x) > 0, ∇f(x) = −λ∇g(x) λ > 0.

Karush-Kuhn-Tucker (KKT) conditions

$$L(\mathbf{x}, \lambda) \equiv f(\mathbf{x}) + \lambda g(\mathbf{x}).$$

$$\begin{aligned} g(\mathbf{x}) &\geqslant 0 \\ \lambda &\geqslant 0 \\ \lambda g(\mathbf{x}) &= 0 \end{aligned}$$



Lagrangian

- Minimize (rather than maximize) the function f(x) subject to an inequality constraint g(x) >= 0

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x})$$

$$L(\mathbf{x}, \{\lambda_j\}, \{\mu_k\}) = f(\mathbf{x}) + \sum_{j=1}^{J} \lambda_j g_j(\mathbf{x}) + \sum_{k=1}^{K} \mu_k h_k(\mathbf{x})$$

-

-