Constraints

Penalty

Barrier

Lagrange

Standard form

minimize $f(\vec{x})$

s.t. $\vec{h}(\vec{x}) = 0$

$\vec{g}(\vec{x}) \leq 0$

Common Types

Bracket

$x_i \in [a, b]$

$x_i \geq a \qquad x_i \leq b$

$0 \geq a - x_i \qquad x_i - b \leq 0$

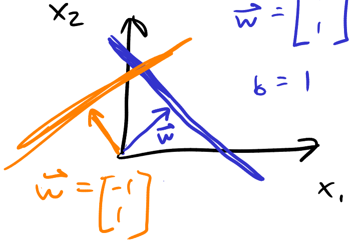
$$\vec{g}(\vec{x}) = \begin{bmatrix} -x_i + a \\ x_i - b \end{bmatrix}$$

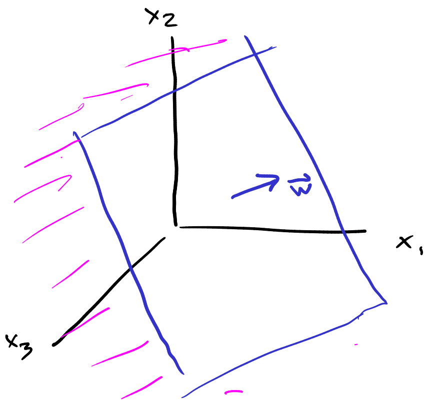
Hyperplane

$\vec{w}^T\vec{x} = b$

"normal vector"

$\vec{h}(\vec{x}) = \vec{w}^T\vec{x} - b$

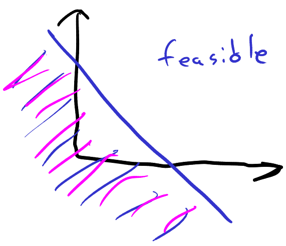




Half-Space

$\vec{w}^T\vec{x} \geq b$

$g(\vec{x}) = -\vec{w}^T\vec{x} + b$



Norm Ball

$\|\vec{x} - \vec{c}\| \leq r$

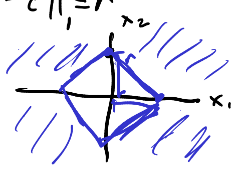
$g(\vec{x}) = \|\vec{x} - \vec{c}\| - r$

$\| \quad \|_2$

$\|\vec{x}\|_1 = \sum_i |x_i|$

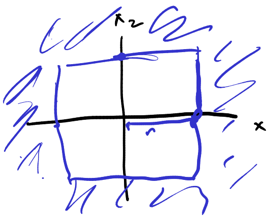
$\| \quad \|_1$

$\|\vec{x} - \vec{c}\|_1 \leq r$



$\max_i x_i$

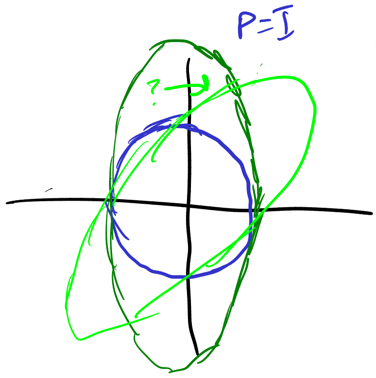
$\| \quad \|_\infty$



Ellipsoid

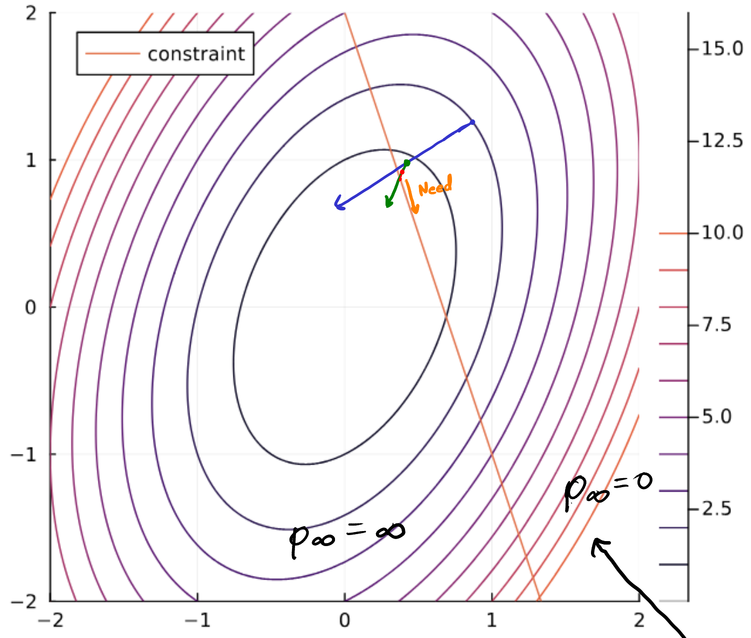
$P \succ 0$

$$(\vec{x}-\vec{c})^T P^{-1}(\vec{x}-\vec{c}) \le 1$$



$P = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$

$P = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 2 \end{bmatrix}$



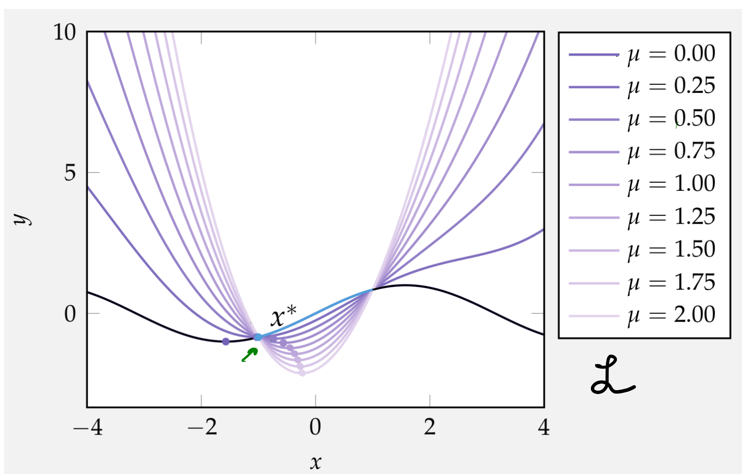
$$p_\infty(\vec{x}) = \begin{cases} 0 & \text{if } g(\vec{x}) \le 0 \\ \infty & \text{o.w.} \end{cases}$$

$$\text{minimize } f(\vec{x}) + p_\infty(\vec{x}) = f(\vec{x}) + \infty \cdot (g(\vec{x}) > 0)$$

equivalent

$$\underset{\vec{x}}{\text{minimize}}\ \underset{\mu \ge 0}{\text{maximize}}\ \underbrace{f(\vec{x}) + \mu g(\vec{x})}_{\mathcal{L}(\vec{x}, \mu)}$$

P1 P2



$\mathcal{L}$

minimize $\sin(x)$
s.t. $x^2 \le 1$

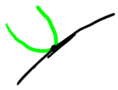
$g(x) = x^2 + 1$

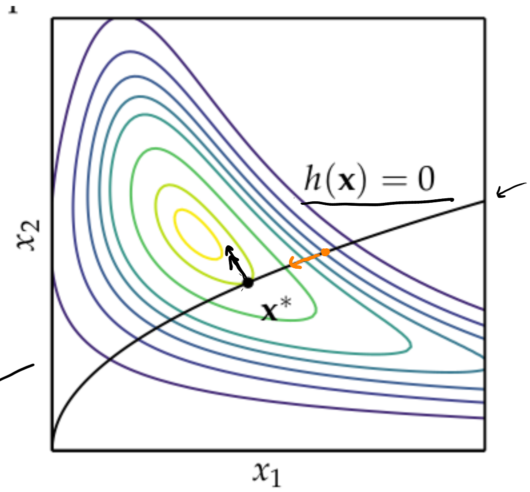
Lagrange Multipliers in 2D for Equality Constraint

$$\underset{x}{\text{minimize}} \quad -\exp\left(-\left(x_1x_2-\tfrac{3}{2}\right)^2-\left(x_2-\tfrac{3}{2}\right)^2\right)$$

subject to $\quad x_1 - x_2^2 = 0$

optima will always occur when constraint satisfaction lines are parallel to contour lines





$$\nabla f(\vec{x})\Big|_{\vec{x}^*} = \lambda \nabla h(\vec{x})\Big|_{\vec{x}^*}$$

$$\mathcal{L}(\vec{x},\lambda) = f(\vec{x}) - \lambda h(\vec{x})$$

$$\nabla \mathcal{L}(\vec{x},\lambda) = 0$$

$\nabla_{\vec{x}} \mathcal{L} = 0$ gives us that $\nabla f = \lambda \nabla h$

$\nabla_{\lambda} \mathcal{L} = 0$ gives us that $h(\vec{x}) = 0$

calculate $\dfrac{\partial \mathcal{L}}{\partial x_1}, \dfrac{\partial \mathcal{L}}{\partial x_2}, \dfrac{\partial \mathcal{L}}{\partial \lambda} = 0$

$x_1 = 1.358$
$x_2 = 1.165$
$\lambda = 0.170$

General Lagrangian

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

$$\underset{\vec{x}}{\text{minimize}} \ \underset{\vec{\mu}\geq 0,\ \vec{\lambda}}{\text{maximize}} \ \mathcal{L}(\vec{x},\vec{\mu},\vec{\lambda})$$

FONC

(unconstrained $\nabla f(x) = 0$)

Karush-Kuhn-Tucker (KKT) conditions

Feasibility

$$\vec{g}(x^*) \leq 0$$
$$\vec{h}(\vec{x}^*) = 0$$

Dual Feasibility

$$\vec{\mu}^* \geq 0$$

Complementary Slackness

$$\vec{\mu}^* \odot \vec{g}(\vec{x}^*) = 0 \qquad (\text{either } \mu_i \text{ or } g_i \text{ is } 0)$$

Stationarity

$$0 = \nabla f(\vec{x}) + \sum_i \lambda_i \nabla h_i(x) + \sum_i \mu_i \nabla g_i(x)$$

$$\nabla_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\mu}, \vec{\lambda}) \Big|_{\vec{x}^*, \vec{\mu}^*, \vec{\lambda}^*} = 0$$

Duality

$$\underset{\vec{x}}{\text{minimize}} \ \underset{\vec{\mu} \geq 0, \vec{\lambda}}{\text{maximize}} \ \mathcal{L}(\vec{x}, \vec{\mu}, \vec{\lambda})$$

"primal problem"

$$\underset{\vec{\mu} \geq 0, \vec{\lambda}}{\text{maximize}} \ \underbrace{\underset{\vec{x}}{\text{minimize}} \ \mathcal{L}(\vec{x}, \vec{\mu}, \vec{\lambda})}$$

"dual problem"

max-min inequality

$$\max_a \min_b f(a,b) \leq \min_b \max_a f(a,b)$$

value of dual problem

$$\boxed{d^* \leq p^*}$$

value of primal problem

"Weak Duality"

"Dual function"

$$D(\vec{\mu}, \vec{\lambda}) \equiv \underset{\vec{x}}{\text{minimize}} \ \mathcal{L}(\vec{x}, \vec{\mu}, \vec{\lambda})$$

$$\max_{\vec{\mu}, \vec{\lambda}} D(\vec{\mu}, \vec{\lambda}) \leq p^*$$

$$D(\vec{\mu}, \vec{\lambda}) \leq p^*$$

If $f(\vec{x}) - D(\vec{\mu}, \vec{\lambda}) \leq \varepsilon$ then $\vec{x}$ is ε-suboptimal

If $\varepsilon = 0$ certificate of optimality

$$\boxed{d^* = p^*}$$

"Strong Duality"

Not true in general

but for convex problems

"usually true"

~Boyd + Vandenberg