* $$\min_{x\in\mathbb{R}^n} f(x)$$
s.t. $h_i(x)\le 0 \quad \forall i\in\{1,2,\dots,m\}$
$\ell_j(x)=0 \quad \forall j\in\{1,2,\dots,r\}$

$$L(x,\lambda,\nu) = f(x) + \sum_{i=1}^{m}\lambda_i h_i(x) + \sum_{j=1}^{r}\nu_j \ell_j(x) \quad ; \lambda_i \ge 0$$
Lagrangian

$$g(\lambda,\nu) = \min_x L(x,\lambda,\nu)$$

Today's lecture $\longrightarrow$ FxTS-GF
↳ Equality constrained optimization problems
↳ Saddle point problems.

Discretization $\longrightarrow$ Augmented Lagrangian Methods / Method of Multipliers

Mechanism of convergence $\longrightarrow$ ADMM (Alternating Direction Method of Multipliers)

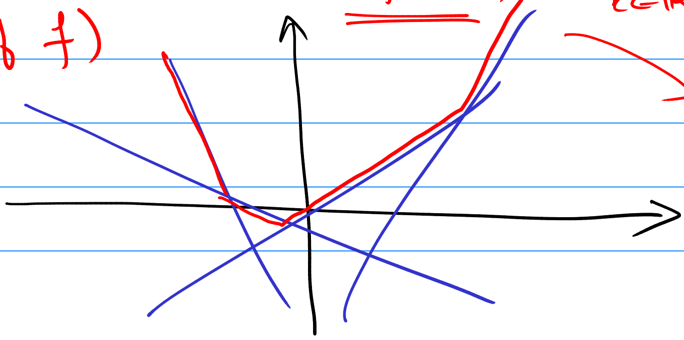
Distributed Optimization $\longleftarrow$ Distributed ADMM

———X——————X——————X——————

FxTS-GF: Unconstrained Optimization { $f(x)$ is SC)
↳
$$\dot{x} = -\frac{\nabla f}{\|\nabla f\|^{\frac{p-2}{p-1}}} - \frac{\nabla f}{\|\nabla f\|^{\frac{q-2}{q-1}}}, \quad p>2, \; q\in(1,2)$$

* Conjugate of a Function f (Fenchel dual of f)

$$f^*(y) := \max_{x\in\mathbb{R}^n}\{y^Tx - f(x)\}$$
always convex !!



* **Theorem:** (i) f is μ-strongly convex $\Rightarrow$ f^* is $\frac{1}{\mu}$-smooth.

(ii) f is L-smooth and convex $\Rightarrow$ f^* is $\frac{1}{L}$-strongly convex.

* Equality constrained optimization problem:

$$\left.\begin{aligned} &\min_{x\in\mathbb{R}^n} \ f(x) \\ &\text{s.t.} \quad Ax=b \end{aligned}\right\} \longrightarrow \text{Primal problem.}$$

$$L(x,\nu) = f(x) + \nu^T(Ax-b)$$

$$\begin{aligned} g(\nu) &= \min_x \{ f(x) + \nu^T(Ax-b) \} \\ &= \min_x \{ f(x) + (A^T\nu)^T x - \nu^T b \} \\ &= -\nu^T b + \min_x \{ (A^T\nu)^T x + f(x) \} \\ &= -\nu^T b - \underbrace{\max_x \{ -f(x) + (-A^T\nu)^T x \}}_{f^*(-A^T\nu)} \end{aligned}$$

$$\boxed{g(\nu) = -\nu^T b - f^*(-A^T\nu)} \longrightarrow \max_\nu g(\nu)$$

FxTS-GF for above problem:

$$\boxed{\dot{\nu} = \frac{\nabla g(\nu)}{\|\nabla g(\nu)\|^{\frac{p-2}{p-1}}} + \frac{\nabla g(\nu)}{\|\nabla g(\nu)\|^{\frac{q-2}{q-1}}}, \quad \begin{array}{l} p>2 \\ q\in(1,2) \end{array}}$$

Ex:

$$\left.\begin{aligned} &\min_{x\in\mathbb{R}^n} \ \frac{1}{2}x^TQx + c^Tx, \quad Q>0 \\ &\text{s.t.} \quad Ax=b \end{aligned}\right\}$$

$$L(x,\nu) = \frac{1}{2}x^TQx + c^Tx + \nu^T(Ax-b)$$

$$g(\nu) = -\nu^Tb - f^*(-A^T\nu) \quad , \text{ where } f(x) := \frac{1}{2}x^TQx + c^Tx$$

$$g(\nu) = \min_x L(x,\nu)$$

$$Qx^* + c + A^T\nu = 0 \qquad \text{or } x^* = -Q^{-1}(c + A^T\nu)$$

$$g(\nu) = \frac{1}{2}\left(-Q^{-1}(c+A^T\nu)\right)^T Q\left(Q^{-1}(c+A^T\nu)\right) - c^TQ^{-1}(c+A^T\nu)$$
$$-\nu^Tb - \nu^TAQ^{-1}(c+A^T\nu)$$
$$= -\nu^Tb - f^*(-A^T\nu)$$

$$f^*(-A^T\nu) = \frac{1}{2}(c+A^T\nu)^TQ^{-1}(c+A^T\nu)$$

$$f^*(y) = \frac{1}{2}(c-y)^TQ^{-1}(c-y)$$

* Saddle point problems:

$$\max_{z\in\mathbb{R}^m} \min_{x\in\mathbb{R}^n} F(x,z)$$

(x^*, z^*) is a saddle point if

$$F(x^*,z) \leq F(x^*,z^*) \leq F(x,z^*) \qquad \forall x,z$$

We want to come up with equivalent GFs for solving saddle-pt problems

$\hookrightarrow$ Working with Lagrangians for equality constrained optimization problems

Two-player games.

* Assumptions: $F(x,z)$ is locally strictly convex-strictly concave in its arguments with

$$\left.\begin{array}{l} \checkmark\ \nabla^2_{xx} F(x,z) > 0 \\ \checkmark\ \nabla^2_{zz} F(x,z) < 0 \end{array}\right\}$$

$$X \triangleq \begin{bmatrix} x \\ z \end{bmatrix} \qquad \nabla F = \begin{bmatrix} \nabla_x F(x,z) \\ -\nabla_z F(x,z) \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = -\left(\nabla^2 F\right)^{-1} \left(\frac{\nabla F}{\|\nabla F\|^{\frac{p-2}{p-1}}} + \frac{\nabla F}{\|\nabla F\|^{\frac{q-2}{q-1}}} \right) \qquad \begin{matrix} p>2 \\ q\in(1,2) \end{matrix}$$

$$\nabla^2 F \triangleq \begin{bmatrix} \nabla^2_{xx} F & \nabla_{xz} F \\ \nabla_{xz} F & \nabla^2_{zz} F \end{bmatrix}$$

Converges in fixed-time

$V = \frac{1}{2} \|\nabla F(x,z)\|^2 \leftarrow$ Lyapunov candidate

$$\dot{V} \leq -c_1 V^{\alpha_1} - c_2 V^{\alpha_2} \qquad \begin{matrix} \alpha\alpha_1 < 1 \\ \alpha_2 > 1 \end{matrix}$$

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

$$\min f(x) + \frac{c}{2} \|h(x)\|^2 \quad \text{s.t.} \quad h(x) = 0 \qquad c>0$$

Same x^*

Ex. $\min -x^2$ s.t. $\sqrt{x^2-4} = 0$ | $\min -x^2 + \frac{c}{2}(x^2-4)$ s.t. $\sqrt{x^2-4}=0$ $\qquad c>2$

* Quadratic penalty makes the original objective function SC (potentially) for large enough c.

* Softer penalty than something like log-barrier → iterates are strictly confined in the interior

$$\min_x f(x) + \frac{1}{\varepsilon} \log\left(\varepsilon - \|h(x)\|\right)$$

$$\left.\begin{array}{ll}\min\limits_{x\in\mathbb{R}^n} & f(x) \\ \text{s.t.} & h(x)=0\end{array}\right\} \quad L(x,\nu) := f(x) + \nu^T h(x)$$

Augmented Lagrangian

$$\left.\begin{array}{ll}\min\limits_{x\in\mathbb{R}^n} & f(x) + \frac{c}{2}\|h(x)\|^2 \\ \text{s.t.} & h(x)=0\end{array}\right\} \quad L_c(x,\nu) := f(x) + \frac{c}{2}\|h(x)\|^2 + \nu^T h(x)$$

$$L_c(x,\nu) = L(x,\nu) + \frac{c}{2}\|h(x)\|^2$$

$\nu \to \nu^*$

$c \to \infty$

$x \to x^*$

x^*, ν^*

When does Augmented Lagrangian work?

Ex:

$$\min_{x_1,x_2} f(x) := \frac{1}{2}(x_1^2 + x_2^2)$$

$$\text{s.t. } x_1 = 1$$

$x^* = (1,0)$, $\nu^* = -1$

$$g(\nu) = \min_x L(x,\nu)$$

$$= \min_x \left\{ \frac{1}{2}(x_1^2 + x_2^2) + \nu(x_1 - 1) \right\}$$

$x_1^* + \nu^* = 0 \quad \nu^* = -1$

$$L_c(x,\nu) = \frac{1}{2}(x_1^2 + x_2^2) + \nu(x_1 - 1) + \frac{c}{2}(x_1 - 1)^2$$

$$x_1^* + \nu + c(x_1^* - 1) = 0 \qquad x_2^* = 0$$

$$\boxed{x_1^* = \frac{c - \nu}{c+1}}$$

$\nu \to \nu^* \quad x_1^k \to 1$

$$x_1^k = \frac{1 - \nu/c}{1 + 1/c}$$

$c \to \infty \quad x_1^k \to 1$

* Convergence Mechanism:

* Take ν closer to ν^*

* If (x^*, ν^*) satisfy 2nd-order sufficiency conditions and c is large enough, then x^* is a strict local minima of $L_c(\cdot, \nu^*)$ corresponding to ν^*.

* 2^{nd}-order sufficiency conditions:

$$\nabla_x L(x^*, \nu^*) = 0 \quad \text{and} \quad \nabla_\nu L(x^*, \nu^*) = 0$$

$$y^T \nabla_{xx}^2 L(x^*, \nu^*) y > 0 \quad \forall y \neq 0 \quad \text{with} \quad \nabla h(x)^T y = 0$$

$\Rightarrow x^*$ is a strict local minima of $L_c(\cdot, \nu^*)$.

* $$L_c(x, \nu) = \underbrace{f(x) + \nu^T h(x)}_{L(x,\nu)} + \frac{c}{2} \|h(x)\|^2$$

$$\nabla_x L_c(x, \nu) = \nabla_x L(x, \nu) + c\, h(x) \cdot \nabla h(x)$$

$$\nabla_x L_c(x, \nu) = \nabla f(x) + \nabla h(x) (\nu + c\, h(x))$$

$$\nabla_{xx}^2 L_c(x, \nu) = \nabla^2 f(x) + \sum_{i=1}^m (\nu_i + c\, h_i(x)) \nabla^2 h_i(x) + c \nabla h(x) \nabla h(x)^T$$

If the 2^{nd}-order sufficiency conditions hold,

$$\nabla_{xx}^2 L_c(x^*, \nu^*) = \nabla^2 f(x^*) + \sum_{i=1}^m \nu_i^* \nabla^2 h_i(x^*) + c \nabla h(x^*) \nabla h(x^*)^T$$

$$= \nabla_{xx}^2 L(x^*, \nu^*) + c \nabla h(x^*) \nabla h(x^*)^T$$

$$y^T \nabla_{xx}^2 L_c(x^*, \nu^*) y > 0 \quad \text{with} \quad \nabla h(x^*)^T y = 0.$$

Lemma: P, Q be two symmetric matrices.
$Q \geq 0$ and $P > 0$ in the null-space of Q, then
$\exists\, \bar{c}$ s.t.
$P + cQ$ is always P.D. $\forall c \geq \bar{c}$