Last Time:

- Line Search (Armijo)
- Equality + Inequality
- KKT Conditions

Today:

- Algorithms for constrained optimization
- Augmented Lagrangian Method
- Quadratic Program
- More on regularization + line search

Inequality-Constrained Minimization:

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

- KKT Conditions:

$$\nabla f + \underbrace{\left(\frac{\partial c}{\partial x}\right)^T \lambda}_{} = 0 \qquad \text{(stationarity)}$$

$$c(x) \leq 0 \qquad \text{(primal feasibility)}$$

$$\lambda \geq 0 \qquad \text{(dual feasibility)}$$

$$\lambda \odot c(x) = 0 \qquad \text{(complementarity)}$$

Hadamard (element-wise) product

This term acts like a penalty. Make sure sign is consistent (makes cost worse for $c(x)$ infeasible)

Optimization Algorithms

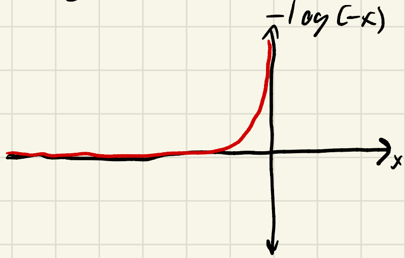
* Active-Set Method

- Guess active/inactive constraints
- Solve equality constrained problem
- Used when you have a good heuristic for active set.

* Barrier/Interior-Point

- Replace inequalities with barrier function in objective:

$$\left.\begin{aligned} &\min_x \; f(x) \\ &\text{s.t. } c(x) \le 0 \end{aligned}\right\} \rightarrow \min_x \; f(x) - \sum_{i=1}^{m} \frac{1}{\rho} \log(-c_i(x))$$

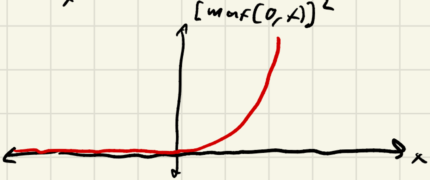


- Gold standard for <u>convex</u> problems

* Penalty

- Replace constraints with penalty term that penalizes violation:

$$\left.\begin{aligned} &\min_x \; f(x) \\ &\text{s.t. } c(x) \le 0 \end{aligned}\right\} \rightarrow \min_x \; f(x) + \frac{\rho}{2} \left[\max(0, c(x))\right]^2$$



- easy to implement
- Has issues with numerical ill-conditioning
- Can't achieve high accuracy

* Augmented Lagrangian
  - Add Lagrange multiplier estimate to penalty method:

$$\min_x \underbrace{f(x) + \tilde{\lambda}^T c(x) + \frac{\rho}{2}[\max(0, c(x))]^2}_{L_\rho(x,\tilde{\lambda}) \text{ "Augmented Lagrangian"}}$$

- Update $\tilde{\lambda}$ by "offloading" penalty term into $\tilde{\lambda}$ at each iteration:

$$\frac{\partial f}{\partial x} + \tilde{\lambda}^T \frac{\partial c}{\partial x} + \rho c(x)^T \frac{\partial c}{\partial x} = \frac{\partial f}{\partial x} + [\tilde{\lambda} + \rho c(x)]^T \frac{\partial c}{\partial x} = 0$$

$$\Rightarrow \tilde{\lambda} \leftarrow \tilde{\lambda} + \rho c(x)$$

(for active constraints)

- Repeat until convergence:
  1) $\min_x L_\rho(x, \tilde{\lambda})$
  2) $\tilde{\lambda} \leftarrow \max(0, \tilde{\lambda} + \rho c(x))$ ← Clamping to guarantee $\tilde{\lambda} \geq 0$
  3) $\rho \leftarrow \alpha \rho$

typically ≈ 10

- Fixes ill-conditioning of penalty method
- Converges with finite ρ
- Works well on non-convex problems

* Quadratic Program

$$\min_x \ \tfrac{1}{2}x^T Q x + q^T x \quad , \quad Q > 0$$
$$\text{s.t.} \quad Ax \leq b$$
$$Cx = d$$

- Super useful in control
- Can be solved very fast ($\sim$ kHz)

* Example

  - Try with penalty, full AL, just λ updates

* Regularization

  - Given:

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

- We might like to turn this into:

$$\min_x \ f(x) + P_\infty(c(x)) \quad , \quad P_\infty(x) = \begin{cases} 0, & c(x) = 0 \\ +\infty, & c(x) \neq 0 \end{cases}$$

- Practically terrible but we can get the same effect by solving:

$$\min_x \max_\lambda \; f(x) + \lambda^T c(x)$$

- Whenever $c(x) \neq 0$, inner max problem blows up
- Similar for inequalities:

$$\min_x f(x) \quad \text{s.t. } c(x) \le 0 \;\; \hookrightarrow \;\; \min_x f(x) + P_\infty^+(c(x))$$

$$P_\infty^+ = \begin{cases} 0, & c(x) \le 0 \\ +\infty, & c(x) > 0 \end{cases}$$

$$\Rightarrow \min_x \max_{\lambda \ge 0} f(x) + \lambda^T c(x)$$

- Interpretation: KKT conditions define a saddle point in (x, λ)
- KKT system should have $\dim(x)$ pos. eigenvalues and $\dim(\lambda)$ neg. eigenvalues at an optimum. Called "quasi definite"

$$\begin{bmatrix} [H + \beta I] & C^T \\ C & [-\beta I] \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x L \\ -c(x) \end{bmatrix}, \quad \beta > 0$$

* Example:

- Still overshoot $\Rightarrow$ need line search