Optimization: Summary

$$\left.\begin{array}{l}\text{Optimization problem}\\ \text{Set of algorithms}\end{array}\right\} \rightarrow \text{most suited algorithm?}$$

1. Simplification of objective function and/or constraints
2. Determination of most efficient available algorithm
3. Determination of stopping criterion

Simplification of objective function or constraint

No systematic procedure.

Examples:

- $$\min_{x \in \mathbb{R}^3} \exp(x_1^2 + x_1 x_2 + 2x_2^2 + x_3^2)$$
  $$\text{s.t.} \quad (x_1 + x_2 + x_3)^2 \leqslant 9$$
  $$\arctan(x_1 - x_2 x_3) \geqslant \frac{\pi}{4}$$

  can be recast as:

  $$\min_{x \in \mathbb{R}^3} x_1^2 + x_1 x_2 + 2x_2^2 + x_3^2$$
  $$\text{s.t.} \quad -3 \leqslant x_1 + x_2 + x_3 \leqslant 3$$
  $$x_1 - x_2 x_3 \geqslant 1$$

Simplification of objective function/constraint (continued)

- $\min_x \|x\|_1 = \min_x \sum_i |x_i| \quad \text{s.t. } Ax \leqslant b$

  can be recast as LP:

  $$\min_\alpha \sum_i \alpha_i \quad \text{s.t.} \quad \begin{aligned} Ax &\leqslant b \\ x &\leqslant \alpha \\ -x &\leqslant \alpha \end{aligned}$$

- $\min_x \|x\|_\infty = \min_x \max_i |x_i| \quad \text{s.t. } Ax \leqslant b$

  can be recast as LP:

  $$\min_t t \quad \text{s.t.} \quad \begin{aligned} Ax &\leqslant b \\ x_i &\leqslant t \\ -x_i &\leqslant t \end{aligned}$$

  Also valid with positive weights added (see lecture notes)

Proof

$$\begin{array}{ccc} (1) & & (2) \\ \min_x \sum_i |x_i| = \min_x \ F_1(x) & \leftrightarrow & \min_\alpha \sum_i \alpha_i = \min \ F_2(\alpha) \\ \text{s.t.} \quad Ax \leqslant b & & \text{s.t.} \quad Ax \leqslant b \\ & & x, -x \leqslant \alpha \end{array}$$

- First prove: feasible solution of (1) $\leftrightarrow$ feasible solution of (2)
- $x^\star$: optimal solution of (1), define $\alpha^\star = |x^\star|$
- $\tilde{x}, \tilde{\alpha}$: optimal solution of (2)

$$\left.\begin{array}{r} \tilde{x} \leqslant \tilde{\alpha} \\ -\tilde{x} \leqslant \tilde{\alpha} \end{array}\right\} \rightarrow |\tilde{x}| \leqslant \tilde{\alpha}$$

Suppose $\tilde{\alpha}_j > |\tilde{x}_j|$ for some j
Set $\tilde{\alpha}_j = |\tilde{x}_j|$
$\Rightarrow$ still feasible, but lower value of F_2 $\Rightarrow$ impossible
$\Rightarrow$ $\tilde{\alpha}_j = |\tilde{x}_j|$ for all j, so $\tilde{\alpha} = |\tilde{x}|$

Proof (continued)

$$\begin{array}{lcl} \text{(1)} & & \text{(2)} \\ \min\limits_x \sum\limits_i |x_i| = \min\limits_x \; F_1(x) & \leftrightarrow & \min\limits_\alpha \sum\limits_i \alpha_i = \min \; F_2(\alpha) \\ \text{s.t. } \; Ax \leqslant b & & \text{s.t. } \; Ax \leqslant b \\ & & \qquad x, -x \leqslant \alpha \end{array}$$

- $x^\star, \alpha^\star$ is feasible solution of problem (2)
  $\Rightarrow F_1(\tilde{x}) = F_2(\tilde{\alpha}) \leqslant F_2(\alpha^\star) = F_1(x^\star)$
  $\Rightarrow \tilde{x}$ is also optimal solution of (1)
- $\tilde{x}$ is feasible solution of problem (1)
  $\Rightarrow F_2(\alpha^\star) = F_1(x^\star) \leqslant F_1(\tilde{x}) = F_2(\tilde{\alpha})$
  $\Rightarrow x^\star, \alpha^\star$ is also optimal solution of (2)

Examples

1. $$\min_{x} \ \max(x_1, x_2, x_3, x_4)$$
   $$\text{s.t.} \ \ |x_1 + x_2 - 2x_3 + x_4 + 9| \leqslant 2$$

2. $$\min_{x} |x_1| + |x_2| + |x_3|$$
   $$\begin{aligned} \text{s.t.} \quad & x_1 + x_2 + x_3 \geqslant 3 \\ & x_1 - 2x_2 + 4x_3 \geqslant 1 \end{aligned}$$

$\rightarrow$ is LP / can be recast as LP?

Selection of algorithm

1. Linear objective + linear constraints:
   simplex (or interior point, if $>$1000 variables or constraints)
2. (Convex) quadratic objective + linear constraints:
   modified simplex (or interior point)
3. Convex objective + convex constraints:
   cutting-plane, ellipsoid, or interior-point
4. Multiple local minima:
   multi-start local optimization, multi-run simulated annealing, or multi-run genetic algorithm
5. Nonlinear non-convex problems
   $\rightarrow$ use gradient and Hessian if possible!
   analytic/numerical computation

Selection of algorithm

- Nonlinear non-convex, unconstrained:
  1. Levenberg-Marquardt or Newton
  2. quasi-Newton algorithm
  3. steepest descent
  4. Powell's perpendicular method (or Nelder-Mead)
- Nonlinear non-convex constrained:
  1. equality constraints:
     elimination, Lagrange method
  2. linear constraints:
     gradient projection or SQP
  3. nonlinear constraints:
     SQP, or penalty or barrier function

Stopping criterion

- Simplex & modified simplex: finite number of steps
- Convex optimization algorithms:

$$|f(x^*) - f(x_k)| \leqslant \varepsilon_f, \quad g(x_k) \leqslant \varepsilon_g$$

$$\|x^* - x_k\|_2 \leqslant \varepsilon_x \qquad \text{(for ellipsoid)}$$

- Unconstrained nonlinear optimization: $\| \nabla f(x_k) \|_2 \leqslant \varepsilon_\nabla$
- Constrained nonlinear optimization:

$$\begin{aligned} \| \nabla f(x_k) + \nabla g(x_k)\,\mu + \nabla h(x_k)\,\lambda \|_2 &\leqslant \varepsilon_{\mathrm{KT},1} \\ | \mu^T g(x_k) | &\leqslant \varepsilon_{\mathrm{KT},2} \\ \mu &\geqslant -\varepsilon_{\mathrm{KT},3} \\ \| h(x_k) \|_2 &\leqslant \varepsilon_{\mathrm{KT},4} \\ g(x_k) &\leqslant \varepsilon_{\mathrm{KT},5} \end{aligned}$$

Stopping criterion (continued)

- Simulated annealing, genetic algorithm, ...:
  maximum number of iterations

  In general, it is recommended to always include this (provided maximum number is taken large enough)

- *Last* resort:

$$\| x_{k+1} - x_k \|_2 \leqslant \varepsilon_x, \quad | f(x_{k+1}) - f(x_k) | \leqslant \varepsilon_f$$

- Combinations

Examples

- $$\max_{x\in\mathbb{R}^3} \; 4x_1 + 5x_2 - 6x_3$$
  $$\text{s.t.} \quad \log|2x_1 + 7x_2 + 5x_3| \leqslant 1$$
  $$x_1, x_2, x_3 \geqslant 0$$

- $$\min_{x\in\mathbb{R}^3} \; \max\left(\cosh(x_1 + x_2 + x_3), (5x_1 - 6x_2 + 7x_3 + 6)^2\right)$$
  $$\text{s.t.} \; \|x\|_2 \leqslant 10$$
  Remark: $\cosh x = \dfrac{e^x + e^{-x}}{2}$

- $$\max_{x\in\mathbb{R}^2} \; e^{-x_1^2 - x_2^2}(x_1^2 + x_1x_2 + 6x_1)$$

- $$\max_{x\in\mathbb{R}^3} \; \frac{x_1x_2x_3}{1 + x_1^6 + x_2^4 + x_3^2}$$
  $$\text{s.t.} \; x_1 + x_2 + x_3 = 1$$

Summary

- $\left.\begin{array}{l}\text{Optimization problem}\\ \text{Set of algorithms}\end{array}\right\} \rightarrow$ most suited algorithm?
- Simplification of objective function and/or constraints
- Determination of most efficient available algorithm
- Stopping criterion