# Primal-Dual Subgradient Method

- equality constrained problems
- inequality constrained problems

# Primal-dual subgradient method

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \qquad i = 1, \ldots, m \\ & Ax = b \end{array}$$

with variable $x \in \mathbf{R}^n$, $f_i : \mathbf{R}^n \to \mathbf{R}$ convex

- primal-dual subgradient method updates both primal and dual variables
- these converge to primal-dual optimal values

# Equality constrained problem

- convex equality constrained problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

with variable $x$ and optimal value $p^\star$

- we will work instead with (equivalent) *augmented problem*

$$\begin{array}{ll} \text{minimize} & f(x) + (\rho/2)\|Ax - b\|_2^2 \\ \text{subject to} & Ax = b \end{array}$$

where $\rho > 0$

# Augmented Lagrangian and optimality conditions

- **augmented Lagrangian** is

$$L(x,\nu) = f(x) + \nu^T(Ax-b) + (\rho/2)\|Ax-b\|_2^2$$

- $(x,\nu)$ primal-dual optimal if and only if

$$\begin{array}{rcl} 0 & \in & \partial_x L(x,\nu) = \partial f(x) + A^T\nu + \rho A^T(Ax-b) \\ 0 & = & -\nabla_\nu L(x,\nu) = b - Ax \end{array}$$

- same as $0 \in T(x,\nu)$, with $z=(x,\nu)$ and $T(x,\nu) = \left[\begin{array}{c} \partial_x L(x,\nu) \\ -\nabla_\nu L(x,\nu) \end{array}\right]$

- $T$ is a **monotone operator** (much more on this later)

# Primal-dual subgradient method

- primal-dual subgradient method is

$$z^{(k+1)} = z^{(k)} - \alpha_k T^{(k)}$$

where $T^{(k)} \in T(z^{(k)})$ and $\alpha_k$ is step length

- more explicitly:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - \alpha_k(g^{(k)} + A^T \nu^{(k)} + \rho A^T(Ax^{(k)} - b)) \\ \nu^{(k+1)} &= \nu^{(k)} + \alpha_k(Ax^{(k)} - b) \end{aligned}$$

where $g^{(k)} \in \partial f(x^{(k)})$

# Convergence

with step size $\alpha_k = \gamma_k / \|T^{(k)}\|_2$,

$$\gamma_k > 0, \quad \sum_k \gamma_k = \infty, \quad \sum_k \gamma_k^2 < \infty$$

we get convergence:

$$f(x^{(k)}) \to p^\star, \qquad Ax^{(k)} - b \to 0$$

# Inequality constrained problem

- convex inequality constrained problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \qquad i = 1, \ldots, m \end{array}$$

  with variable $x$, optimal value $p^\star$

- (equivalent) augmented problem

$$\begin{array}{ll} \text{minimize} & f_0(x) + (\rho/2)\|F(x)\|_2^2 \\ \text{subject to} & F(x) \preceq 0 \end{array}$$

  where $F(x) = (f_1(x)_+, \ldots, f_m(x)_+)$, $\rho > 0$

# Augmented Lagrangian and optimality conditions

- augmented Lagrangian is

$$L(x,\lambda) = f_0(x) + \lambda^T F(x) + (\rho/2)\|F(x)\|_2^2$$

- $(x,\lambda)$ primal-dual optimal if and only if

$$\begin{aligned} 0 &\in \partial_x L(x,\lambda) = \partial f_0(x) + \sum_{i=1}^m (\lambda_i + \rho f_i(x)_+)\partial f_i(x)_+ \\ 0 &= -\nabla_\lambda L(x,\lambda) = -F(x) \end{aligned}$$

# Primal-dual subgradient method

- define $z = (x, \nu)$ and

$$T(x, \lambda) = \left[ \begin{array}{c} \partial_x L(x, \lambda) \\ -\nabla_\lambda L(x, \lambda) \end{array} \right]$$

($T$ is the KKT operator for the problem, and is monotone)

- primal-dual subgradient method is

$$z^{(k+1)} = z^{(k)} - \alpha_k T^{(k)}$$

where $T^{(k)} \in T(z^{(k)})$ and $\alpha_k$ is step length

- more explicitly:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - \alpha_k \left( g_0^{(k)} + \sum_{i=1}^m (\lambda_i^{(k)} + \rho f_i(x^{(k)})_+) g_i^{(k)} \right) \\ \lambda_i^{(k+1)} &= \lambda_i^{(k)} + \alpha_k f_i(x^{(k)})_+, \quad i = 1, \ldots, m \end{aligned}$$

where $g_0^{(k)} \in \partial f_0(x^{(k)})$, $g_i^{(k)} \in \partial f_i(x^{(k)})_+$, $i = 1, \ldots, m$

- note that $\lambda_i^{(k)}$ can only increase with $k$

# Convergence

with step size $\alpha_k = \gamma_k / \|T^{(k)}\|_2$,

$$\gamma_k > 0, \quad \sum_k \gamma_k = \infty, \quad \sum_k \gamma_k^2 < \infty$$

we get convergence:

$$f_0(x^{(k)}) \to p^\star, \qquad f_i(x^{(k)})_+ \to 0, \quad i = 1, \ldots, m$$

# Example: Inequality constrained LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

primal-dual subgradient update is

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - \alpha_k \left( c + A^T M^{(k)} (\lambda^{(k)} + \rho(Ax^{(k)} - b)_+) \right) \\ \lambda^{(k+1)} &= \lambda^{(k)} + \alpha_k (Ax^{(k)} - b)_+ \end{aligned}$$

where $M^{(k)}$ is a diagonal matrix

$$M_{ii}^{(k)} = \begin{cases} 1 & a_i^T x^{(k)} > b_i \\ 0 & a_i^T x^{(k)} \leq b_i \end{cases}$$

problem instance with $n = 20$, $m = 200$, $p^\star \approx -3.4$
step size $\alpha_k = 1/(k\|T^{(k)}\|_2)$