Primal-dual augmented-Lagrangian barrier solver for non-convex optimization

Taylor Howell

May 3, 2020

Background

Optimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0, \\ & x_L \leq x \leq x_U \end{aligned}$$

Barrier problem

$$\begin{aligned} \min_x \quad & \varphi(x;\mu) \coloneqq f(x) - \mu \sum_{i \in I_L} \log(x^{(i)} - x_L^{(i)}) - \mu \sum_{i \in I_U} \log(x_U^{(i)} - x^{(i)}) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned}$$

Formulation

Problem with split constraint set

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & c_{\mathcal{I}}(x) \geq 0, \\
& c_{\mathcal{E}}(x) = 0, \\
& c_{\mathcal{A}}(x) = 0, \\
& x_L \leq x \leq x_U
\end{aligned}
$$

Slack reformulation for inequality constraints

$$
\begin{aligned}
\min_{x,\, s} \quad & f(x) \\
\text{s.t.} \quad & c_{\mathcal{I}}(x) - s = 0, \\
& c_{\mathcal{E}}(x) = 0, \\
& c_{\mathcal{A}}(x) = 0, \\
& x_L \leq x \leq x_U, \\
& s \geq 0
\end{aligned}
$$

Barrier problem

$$
\begin{aligned}
\min_{x,\, s} \quad & \varphi(x, s; \mu) \\
\text{s.t.} \quad & c_{\mathcal{I}}(x) - s = 0, \\
& c_{\mathcal{E}}(x) = 0, \\
& c_{\mathcal{A}}(x) = 0, \\
& (x_L \leq x \leq x_U,\ s \geq 0)
\end{aligned}
$$

$$
\varphi(x, s; \mu) \coloneqq f(x) - \mu \sum_{i \in I_L} \log(x^{(i)} - x_L^{(i)}) - \mu \sum_{i \in I_U} \log(x_U^{(i)} - x^{(i)}) - \mu \sum_{i \in \mathcal{I}} \log(s^{(i)})
$$

Barrier problem with damping

$$
\begin{aligned}
\min_{x,\, s} \quad & \varphi(x, s; \mu) \\
\text{s.t.} \quad & c_{\mathcal{I}}(x) - s = 0, \\
& c_{\mathcal{E}}(x) = 0, \\
& c_{\mathcal{A}}(x) = 0, \\
& (x_L \leq x \leq x_U,\ s \geq 0)
\end{aligned}
$$

$$
\begin{aligned}
\varphi(x, s; \mu) \coloneqq f(x) &- \mu \sum_{i \in I_L} \log(x^{(i)} - x_L^{(i)}) - \mu \sum_{i \in I_U} \log(x_U^{(i)} - x^{(i)}) - \mu \sum_{i \in \mathcal{I}} \log(s^{(i)}) \\
&+ \mu \kappa_D \sum_{i \in I_L \setminus I_U} (x^{(i)} - x_L^{(i)}) + \mu \kappa_D \sum_{i \in I_U \setminus I_L} (x_U^{(i)} - x^{(i)}) + \mu \kappa_D \sum_{i \in \mathcal{I}} s^{(i)}
\end{aligned}
$$

augmented-Lagrangian barrier problem with slack reformulation

$$\begin{aligned}
\min_{x,\,s,\,r} \quad & \varphi_A(x,s,r;\mu,\lambda,\rho) \\
\text{s.t.} \quad & c_{\mathcal{I}}(x) - s = 0, \\
& c_{\mathcal{E}}(x) = 0, \\
& c_{\mathcal{A}}(x) - r = 0, \\
& (x_L \le x \le x_U,\ s \ge 0)
\end{aligned}$$

$$\begin{aligned}
\varphi_{\mathcal{A}}(x,s,r;\mu,\lambda,\rho) \coloneqq f(x) & - \mu \sum_{i\in I_L} \log(x^{(i)} - x_L^{(i)}) - \mu \sum_{i\in I_U} \log(x_U^{(i)} - x^{(i)}) - \mu \sum_{i\in\mathcal{I}} \log(s^{(i)}) \\
& + \mu\kappa_D \sum_{i\in I_L\setminus I_U} (x^{(i)} - x_L^{(i)}) + \mu\kappa_D \sum_{i\in I_U\setminus I_L} (x_U^{(i)} - x^{(i)}) + \mu\kappa_D \sum_{i\in\mathcal{I}} s^{(i)} \\
& + \lambda^T r + \frac{\rho}{2} r^T r
\end{aligned}$$

Primal-dual equations

$$\begin{aligned}
\nabla f(x) + \nabla c_{\mathcal{I}}(x)^T y_{\mathcal{I}} + \nabla c_{\mathcal{E}}(x)^T y_{\mathcal{E}} + \nabla c_{\mathcal{A}}(x)^T y_{\mathcal{A}} - z_L + z_U + \kappa_D \mu e_L - \kappa_D \mu e_U = r_x &= 0 \\
-y_{\mathcal{I}} - z_s + \kappa_D \mu e_s = r_s &= 0 \\
r + \frac{1}{\rho}(\lambda - y_{\mathcal{A}}) = r_r &= 0 \\
c_{\mathcal{I}}(x) - s = r_{y_{\mathcal{I}}} &= 0 \\
c_{\mathcal{E}}(x) = r_{y_{\mathcal{E}}} &= 0 \\
c_{\mathcal{A}}(x) - r = r_{y_{\mathcal{A}}} &= 0 \\
X_L z_L - \mu e = r_{z_L} &= 0 \\
X_U z_U - \mu e = r_{z_U} &= 0 \\
S z_s - \mu e = r_{z_s} &= 0 \\
(x - x_L \ge 0,\ x_U - x \ge 0,\ s \ge 0,\ z_L \ge 0,\ z_U \ge 0)
\end{aligned}$$

KKT system (full-space)

$$\begin{bmatrix}
W & \nabla c_{\mathcal{I}}^T & \nabla c_{\mathcal{E}}^T & \nabla c_{\mathcal{A}}^T & -I & I & 0 & 0 & 0 \\
\nabla c_{\mathcal{I}} & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 \\
\nabla c_{\mathcal{E}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\nabla c_{\mathcal{A}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I \\
Z_L & 0 & 0 & 0 & X_L & 0 & 0 & 0 & 0 \\
-Z_U & 0 & 0 & 0 & 0 & X_U & 0 & 0 & 0 \\
0 & -I & 0 & 0 & 0 & 0 & 0 & -I & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & Zs & S & 0 \\
0 & 0 & 0 & -\frac{1}{\rho}I & 0 & 0 & 0 & 0 & I
\end{bmatrix}
\begin{bmatrix} dx \\ dy_{\mathcal{I}} \\ dy_{\mathcal{E}} \\ dy_{\mathcal{A}} \\ dz_L \\ dz_U \\ ds \\ dz_s \\ dr \end{bmatrix}
= -
\begin{bmatrix} r_x \\ r_{y_{\mathcal{I}}} \\ r_{y_{\mathcal{E}}} \\ r_{y_{\mathcal{A}}} \\ r_{z_L} \\ r_{z_U} \\ r_s \\ r_{z_s} \\ r_r \end{bmatrix}$$

$$W(x, y_{\mathcal{I}}, y_{\mathcal{E}}, y_{\mathcal{A}}) = \nabla^2 f(x) + \textstyle\sum_{i=1}^{m_{\mathcal{I}}} y_{\mathcal{I}}^{(i)} \nabla^2 c_{\mathcal{I}}(x) + \sum_{i=1}^{m_{\mathcal{E}}} y_{\mathcal{E}}^{(i)} \nabla^2 c_{\mathcal{E}}(x) + \sum_{i=1}^{m_{\mathcal{A}}} y_{\mathcal{A}}^{(i)} \nabla^2 c_{\mathcal{A}}(x)$$

symmetric KKT system (reduced-space)

$$\begin{bmatrix}
\bar{W} & \nabla c_{\mathcal{I}}^T & \nabla c_{\mathcal{E}}^T & \nabla c_{\mathcal{A}}^T \\
\nabla c_{\mathcal{I}} & -Z_s^{-1} S & 0 & 0 \\
\nabla c_{\mathcal{E}} & 0 & 0 & 0 \\
\nabla c_{\mathcal{A}} & 0 & 0 & -\frac{1}{\rho} I
\end{bmatrix}
\begin{bmatrix} dx \\ dy_{\mathcal{I}} \\ dy_{\mathcal{E}} \\ dy_{\mathcal{A}} \end{bmatrix}
= -
\begin{bmatrix} r_x + X_L^{-1} r_{z_L} - X_U^{-1} r_{z_U} \\ r_{y_{\mathcal{I}}} + Z_s^{-1}(S r_s + r_{z_s}) \\ r_{y_{\mathcal{E}}} \\ r_{y_{\mathcal{A}}} + r_r \end{bmatrix}$$

$$\bar{W}(x, y_{\mathcal{I}}, y_{\mathcal{E}}, y_{\mathcal{A}}) = W(x, y_{\mathcal{I}}, y_{\mathcal{E}}, y_{\mathcal{A}}) + X_L^{-1} Z_L + X_U^{-1} Z_U$$

$$\begin{aligned}
dr &= \frac{1}{\rho} dy_{\mathcal{A}} - r_r \\
dz_L &= -X_L^{-1} Z_L dx - X_L^{-1} r_{z_L} \\
dz_U &= X_U^{-1} Z_U dx - X_U^{-1} r_{z_U} \\
dz_s &= -dy_{\mathcal{I}} + r_s \\
ds &= -Z_s^{-1} S dz_s - Z_s^{-1} r_{z_s}
\end{aligned}$$

Restoration phase

Formulation

KKT system (full-space)

$$
\begin{bmatrix}
W & \nabla c_{\mathcal{I}}^T & \nabla c_{\mathcal{E}}^T & \nabla c_{\mathcal{A}}^T & -I & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\nabla c_{\mathcal{I}} & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & -I & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\nabla c_{\mathcal{E}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\nabla c_{\mathcal{A}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & -I & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\
Z_L & 0 & 0 & 0 & X_L & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-Z_U & 0 & 0 & 0 & 0 & X_U & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -I & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & Zs & S & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{\rho}I & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 \\
0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{p_{\mathcal{I}}} & 0 & 0 & 0 & 0 & 0 & P_{\mathcal{I}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{p_{\mathcal{E}}} & 0 & 0 & 0 & 0 & 0 & P_{\mathcal{E}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{p_{\mathcal{A}}} & 0 & 0 & 0 & 0 & 0 & P_{\mathcal{A}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{n_{\mathcal{I}}} & 0 & 0 & 0 & 0 & 0 & N_{\mathcal{I}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{n_{\mathcal{E}}} & 0 & 0 & 0 & 0 & 0 & N_{\mathcal{E}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_{n_{\mathcal{A}}} & 0 & 0 & 0 & 0 & 0 & N_{\mathcal{A}}
\end{bmatrix}
\begin{bmatrix}
d \\ dy \\ dy \\ dy \\ dz \\ dz \\ d \\ dz \\ d \\ dp \\ dp \\ dp \\ dn \\ dn \\ dn \\ dz \\ dz \\ dz \\ dz \\ dz \\ dz
\end{bmatrix}
$$

symmetric KKT system (reduced-space)

$$
\begin{bmatrix}
\bar{W} & \nabla c_{\mathcal{I}}^T & \nabla c_{\mathcal{E}}^T & \nabla c_{\mathcal{A}}^T \\
\nabla c_{\mathcal{I}} & \Sigma_{\mathcal{I}} & 0 & 0 \\
\nabla c_{\mathcal{E}} & 0 & \Sigma_{\mathcal{E}} & 0 \\
\nabla c_{\mathcal{A}} & 0 & 0 & \Sigma_{\mathcal{A}}
\end{bmatrix}
\begin{bmatrix}
dx \\ dy_{\mathcal{I}} \\ dy_{\mathcal{E}} \\ dy_{\mathcal{A}}
\end{bmatrix}
= -
\begin{bmatrix}
r_x + X_L^{-1} r_{z_L} - X_U^{-1} r_{z_U} \\
r_{y_{\mathcal{I}}} + Z_s^{-1}(Sr_s + r_{z_s}) + Z_{p_{\mathcal{I}}}^{-1}(P_{\mathcal{I}} r_{p_{\mathcal{I}}} + r_{zp_{\mathcal{I}}}) - Z_{n_{\mathcal{I}}}^{-1}(N_{\mathcal{I}} r_{n_{\mathcal{I}}} + r_{zn_{\mathcal{I}}}) \\
r_{y_{\mathcal{E}}} + Z_{p_{\mathcal{E}}}^{-1}(P_{\mathcal{E}} r_{p_{\mathcal{E}}} + r_{zp_{\mathcal{E}}}) - Z_{n_{\mathcal{E}}}^{-1}(N_{\mathcal{E}} r_{n_{\mathcal{E}}} + r_{zn_{\mathcal{E}}}) \\
r_{y_{\mathcal{A}}} + r_r + Z_{p_{\mathcal{A}}}^{-1}(P_{\mathcal{A}} r_{p_{\mathcal{A}}} + r_{zp_{\mathcal{A}}}) - Z_{n_{\mathcal{A}}}^{-1}(N_{\mathcal{A}} r_{n_{\mathcal{A}}} + r_{zn_{\mathcal{A}}})
\end{bmatrix}
$$

$$
\begin{aligned}
\Sigma_{\mathcal{I}} &= -(Z_s^{-1} s + Z_{p_{\mathcal{I}}}^{-1} p_{\mathcal{I}} - Z_{n_{\mathcal{I}}}^{-1} n_{\mathcal{I}}) \\
\Sigma_{\mathcal{E}} &= -(Z_{p_{\mathcal{E}}}^{-1} p_{\mathcal{E}} - Z_{n_{\mathcal{E}}}^{-1} n_{\mathcal{E}}) \\
\Sigma_{\mathcal{A}} &= -(Z_{p_{\mathcal{A}}}^{-1} p_{\mathcal{A}} - Z_{n_{\mathcal{A}}}^{-1} n_{\mathcal{A}} + \frac{1}{\rho} I)
\end{aligned}
$$

$$
\begin{aligned}
dr &= \frac{1}{\rho} dy_{\mathcal{A}} - r_r \\
dz_L &= -X_L^{-1} Z_L dx - X_L^{-1} r_{z_L} \\
dz_U &= X_U^{-1} Z_U dx - X_U^{-1} r_{z_U} \\
dz_s &= -dy_{\mathcal{I}} + r_s \\
ds &= -Z_s^{-1} S dz_s - Z_s^{-1} r_{z_s} \\
dz_p &= -dy + r_p \\
dz_n &= dy + r_n \\
dp &= Z_p^{-1}(-P dz_p - r_{z_p}) \\
dn &= Z_n^{-1}(-N dz_n - r_{z_n})
\end{aligned}
$$