Implementation and Solution of IPOPT Primal-Dual System

Victor M. Zavala

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

This document provides some details on the implementation and solution of the primal-dual system in IPOPT [1]. The document is expected to be a quick reference for the development and implementation of large-scale computational strategies.

Consider the nonlinear programming problem of the form,

$$\begin{aligned} \min \quad & f(x) \\ & g_L \le g(x) \le g_U \\ & x_L \le x \le x_U \end{aligned} \tag{1}$$

where $x \in \Re^n$ are the primal variables with lower and upper bounds $x_L \in \Re^n$, $x_U \in \Re^n$. The inequality constraints $g : \Re^n \to \Re^m$ are bounded by $g_L \in \Re^m$ and $g_U \in \Re^m$.

After this problem has been communicated to IPOPT, the solver makes an explicit distinction between the equality (defined with $g_L = g_U$) and inequality constraints to give,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & d_L \le d(x) \le d_U \\ & x_L \le x \le x_U \end{aligned} \tag{2}$$

The equality constraints are represented by $c : \Re^n \to \Re^{m_c}$ and $d : \Re^n \to \Re^{m_d}$ denotes the inequality constraints with bounds $d_L \in \Re^{m_d}$ and $d_U \in \Re^{m_d}$ and $m = m_c + m_d$. Having done this, the current implementation of IPOPT reformulates the general inequality constraints by adding slack variables and their corresponding bounds,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & d(x) - s = 0 \\ & x - x_L \ge 0, \quad x_U - x \ge 0 \\ & s - d_L \ge 0, \quad d_U - s \ge 0 \end{aligned} \tag{3}$$

with $s \in \Re^{n_d}$. As required by IPOPT, if a variable bound does not exist, the user sets the corresponding value to a large number ($-\infty$ or ∞). Nevertheless, for efficiency reasons, the solver ensures that only the relevant specified bounds ($x_L, d_L > -\infty$ and $x_U, d_U < \infty$) are actually taken into account. This is done by reformulating the problem to,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & d(x) - s = 0 \\ & (P_x^L)^T x - x_L \ge 0, \quad x_U - (P_x^U)^T x \ge 0 \\ & (P_d^L)^T d(x) - d_L \ge 0, \quad d_U - (P_d^U)^T d(x) \ge 0 \end{aligned} \tag{4}$$

where $P_x^L \in \Re^{n \times n_{xL}}$, $P_x^U \in \Re^{n \times n_{xU}}$, $P_d^L \in \Re^{m_d \times n_{dL}}$ and $P_d^U \in \Re^{m_d \times n_{dU}}$ are projection or permutation matrices between variables x and the inequalities $d(x)$ and their corresponding bounds. Symbols n_{xL}, n_{xU}, n_{dL} and n_{dU} represent the number of valid bounds. Accordingly, notice that the dimensions of x_L, x_U, d_L and d_U are also reduced.

In order to derive the primal-dual system, we define the Lagrange function of the reformulated NLP (4) as,

$$\begin{aligned} \mathcal{L} = f(x) + y_c^T c(x) + y_d^T (d(x) - s) - z_L^T \left((P_x^L)^T x - x_L\right) - z_U^T \left(x_U - (P_x^U)^T x\right) \\ -\nu_L^T \left((P_d^L)^T d(x) - d_L\right) - \nu_U^T \left(d_U - (P_d^U)^T d(x)\right) \end{aligned} \tag{5}$$

where $y_c \in \Re^{m_c}$ and $y_d \in \Re^{m_d}$ are the Lagrange multipliers for the equality and inequality constraints, respectively; $z_L \in \Re^{n_{xL}}$ and $z_U \in \Re^{n_{xU}}$ are multipliers for the lower and upper bounds of the x variables; and $\nu_L \in \Re^{n_{dL}}$ and $\nu_U \in \Re^{n_{dU}}$ are the bound multipliers corresponding to the slack variables (multipliers of inequality constraints).

After eliminating the bounds by adding a logarithmic barrier term to the objective function, the primal-dual optimality conditions of problem (4) are given by:

$$\begin{aligned} \nabla_x \mathcal{L} = \nabla_x f(x) + J_c(x)^T y_c + J_d(x)^T y_d - P_x^L z_L + P_x^U z_U &= 0 \\ \nabla_s \mathcal{L} = -y_d - P_d^L \nu_L + P_d^U \nu_U &= 0 \\ Sl_x^L Z_L e - \mu e &= 0 \\ Sl_x^U Z_U e - \mu e &= 0 \\ Sl_d^L V_L e - \mu e &= 0 \\ Sl_d^U V_U e - \mu e &= 0 \\ c(x) &= 0 \\ d(x) - s &= 0 \end{aligned} \tag{6}$$

where $J_c^T \in \Re^{n \times m_c}$ and $J_d^T \in \Re^{n \times m_d}$ are the Jacobian matrices of the equality and inequality constraints and the diagonal matrices,

$$\begin{aligned} Z_L &= \mathrm{diag}(z_L) \\ Sl_x^L &= \mathrm{diag}\left((P_x^L)^T x - x_L\right) \\ Z_U &= \mathrm{diag}(z_U) \\ Sl_x^U &= \mathrm{diag}\left(x_U - (P_x^U)^T x\right) \\ V_L &= \mathrm{diag}(\nu_L) \\ Sl_d^L &= \mathrm{diag}\left((P_d^L)^T d(x) - d_L\right) \\ V_U &= \mathrm{diag}(\nu_U) \\ Sl_d^U &= \mathrm{diag}\left(d_U - (P_d^U)^T d(x)\right) \end{aligned} \tag{7}$$

have appropriate dimensions.

The optimality conditions (6) can be viewed as a set of nonlinear equations parameterized in the scalar parameter μ. For the solution of this system, we can derive a sequence of Newton steps obtained from the linearization of the above expressions,

$$W \Delta x + J_c^T \Delta y_c + J_d^T \Delta y_d - P_x^L \Delta z_L + P_x^U \Delta z_U = -\nabla_x \mathcal{L}$$

$$
\begin{aligned}
-\Delta y_d - P_d^L \Delta\nu_L + P_d^U \Delta\nu_U &= -\nabla_s \mathcal{L} \\
Z_L (P_x^L)^T \Delta x + Sl_x^L \Delta z_L &= -(Sl_x^L Z_L e - \mu e) \\
-Z_U (P_x^U)^T \Delta x + Sl_x^U \Delta z_U &= -(Sl_x^U Z_U e - \mu e) \\
V_L (P_d^L)^T \Delta s + Sl_d^L \Delta\nu_L &= -(Sl_d^L V_L e - \mu e) \\
-V_U (P_d^U)^T \Delta s + Sl_d^U \Delta\nu_U &= -(Sl_d^U V_U e - \mu e) \\
J_c \Delta x &= -c(x) \\
J_d \Delta x - \Delta s &= -(d(x) - s)
\end{aligned}
\tag{8}
$$

where $W \in \Re^{n\times n}$ is the Hessian matrix. The system of linear equations (8) has the following structure,

$$
\begin{bmatrix}
W & 0 & J_c^T & J_d^T & -P_x^L & P_x^U & 0 & 0 \\
0 & 0 & 0 & -I & 0 & 0 & -P_d^L & P_d^U \\
J_c & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
J_d & -I & 0 & 0 & 0 & 0 & 0 & 0 \\
Z_L(P_x^L)^T & 0 & 0 & 0 & Sl_L^x & 0 & 0 & 0 \\
-Z_U(P_x^U)^T & 0 & 0 & 0 & 0 & Sl_U^x & 0 & 0 \\
0 & V_L(P_d^L)^T & 0 & 0 & 0 & 0 & Sl_L^s & 0 \\
0 & -V_U(P_d^U)^T & 0 & 0 & 0 & 0 & 0 & Sl_U^s
\end{bmatrix}
\begin{pmatrix}
\Delta x \\ \Delta s \\ \Delta y_c \\ \Delta y_d \\ \Delta z_L \\ \Delta z_U \\ \Delta v_L \\ \Delta v_U
\end{pmatrix}
= -
\begin{pmatrix}
\nabla_x \mathcal{L} \\ \nabla_s \mathcal{L} \\ c(x) \\ d(x) - s \\ Sl_x^L Z_L e - \mu e \\ Sl_x^U Z_U e - \mu e \\ Sl_d^L V_L e - \mu e \\ Sl_d^U V_U e - \mu e
\end{pmatrix}
\tag{9}
$$

we will refer to this set of linear equations as the primal-dual system. The solution of this system is usually the most expensive step in the algorithm. In the current implementation of IPOPT, the primal-dual system is decomposed by eliminating the bound multipliers leading to the augmented linear system,

$$
\begin{bmatrix}
W + D_x & 0 & J_c^T & J_d^T \\
0 & D_s & 0 & -I \\
J_c & 0 & 0 & 0 \\
J_d & -I & 0 & 0
\end{bmatrix}
\begin{pmatrix}
\Delta x \\ \Delta s \\ \Delta y_c \\ \Delta y_d
\end{pmatrix}
= -
\begin{pmatrix}
\nabla_x \bar{\mathcal{L}} \\ \nabla_s \bar{\mathcal{L}} \\ c(x) \\ d(x) - s
\end{pmatrix}
\tag{10}
$$

where,

$$
\begin{aligned}
\nabla_x \bar{\mathcal{L}} &= \nabla_x f(x) + J_c^T y_c + J_d^T y_d + P_x^U (Sl_x^U)^{-1} \mu e - P_x^L (Sl_x^L)^{-1} \mu e \\
\nabla_s \bar{\mathcal{L}} &= -y_d + P_d^U (Sl_d^U)^{-1} \mu e - P_d^L (Sl_d^L)^{-1} \mu e \\
D_x &= P_x^L (Sl_x^L)^{-1} Z_L (P_x^L)^T - P_x^U (Sl_x^U)^{-1} Z_U (P_x^U)^T \\
D_s &= P_d^L (Sl_d^L)^{-1} V_L (P_d^L)^T - P_d^U (Sl_d^U)^{-1} V_U (P_d^U)^T.
\end{aligned}
$$

Once the augmented linear system is solved, we can obtain step directions for the bound multipliers from,

$$
\begin{aligned}
\Delta z_L &= -z_L + (Sl_x^L)^{-1} \left(\mu e - Z_L (P_x^L)^T \Delta x\right) \\
\Delta z_U &= z_U + (Sl_x^U)^{-1} \left(\mu e - Z_U (P_x^U)^T \Delta x\right) \\
\Delta \nu_L &= -\nu_L + (Sl_d^L)^{-1} \left(\mu e - V_L (P_d^L)^T \Delta s\right) \\
\Delta \nu_U &= \nu_U + (Sl_d^U)^{-1} \left(\mu e - V_U (P_d^U)^T \Delta s\right)
\end{aligned}
\tag{11}
$$

References

[1] Wächter, A. and Biegler, L.T. On The Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Math. Programm.*, **2006**, *106*, 25-57.