

EIN NUMERISCH STABILER CONDENSING ALGORITHMUS FÜR OCP

ROBERT SCHOLZ

1. EINLEITUNG

Das Multiple Shooting Verfahren in Kombination mit sequentieller quadratische Programmierung (SQP) ist eine gängige Methode zur Lösung von Optimale Steuerung Problemen (OCP). Die resultierenden quadratischen Subprobleme sind in der Regel dünn besetzt und haben eine charakteristische Struktur. Sogenannte Condensing Algorithmen verringern den Rechenaufwand, indem sie diese Struktur ausnutzen. Dabei wird das ursprüngliche dünnbesetzte QP in ein dicht besetztes QP mit kleinerer Dimension transformiert. Die klassischen Ansätze bringen jedoch numerische Schwierigkeiten mit sich, wenn das Problem schlecht konditioniert ist. In diesem Bericht stellen wir einen Condensing Algorithmus vor, der auf QR-Faktorisierung beruht. Da nur mit orthogonalen Matrizen gearbeitet wird, zeichnet dieser Ansatz sich durch numerische Stabilität aus.

Wir behandeln quadratische Programme mit der folgenden Struktur. Dabei gibt n_u die Anzahl der Steuerung, n_s die Anzahl der Zustände, n_p die Anzahl der Parameter und N die Anzahl der Shootingknoten an. Es ist also $u_i \in \mathbb{R}^{n_u}$ und $s_i \in \mathbb{R}^{n_s}$ für $i = 0, \dots, N$ sowie $p \in \mathbb{R}^{n_p}$. Um die Darstellung einfach zu halten, gehen wir davon aus, dass die gekoppelten sowie die ungekoppelten Nebenbedingungen mit Gleichheit erfüllt sein müssen. Der Ansatz lässt sich jedoch direkt auf Ungleichungsnebenbedingungen erweitern.

$$\begin{aligned}
 \min_{u,s,p} \quad & p^\top B_{pp} p + f_p^\top p \\
 & + \frac{1}{2} \sum_{i=1}^N s_i^\top B_{ss}^i s_i + u_i^\top B_{uu}^i u_i \\
 & + \sum_{i=1}^N s_i^\top B_{su}^i u_i + p^\top B_{ps}^i s_i + p^\top B_{pu}^i u_i \\
 & + \sum_{i=1}^N (f_u^i)^\top u_i + (f_s^i)^\top s_i \\
 \text{s.t.} \quad & s_{i+1} = X_p^i p + X_s^i s_i + X_u^i u_i - X_c^i \quad i = 1, \dots, N-1 \\
 & R_u^i u_i + R_s^i s_i + R_p^i p = R_c \quad i = 1, \dots, N \\
 & \sum_{i=1}^N C_u^i u_i + C_s^i s_i + R_p^i p = C_c \\
 & p_l \leq p \leq p_u \\
 & s_l^i \leq s_i \leq s_u^i \quad i = 1, \dots, N \\
 & u_l^i \leq u_i \leq u_u^i \quad i = 1, \dots, N
 \end{aligned}$$

2. DER ALGORITHMUS

2.1. Anordnung der einzelnen Matrizen. Um die Notation einfach zu halten, fassen wir die Zielfunktion und Nebenbedingungen in einzelne Matrizen zusammen. Dabei bleibt die Struktur erhalten. Wir definieren $x \in \mathbb{R}^{n_p+N(n_u+n_s)}$ mit $x = (p, s_0, u_0, s_1, u_1, \dots, s_N, u_N)$. Somit können wir das QP auch schreiben als:

$$(2.1) \quad \begin{aligned} x^* = \min_x \quad & \frac{1}{2} x^\top H x + f^\top x \\ \text{s.t.} \quad & Sx = s \\ & Qx = q \\ & b_l \leq x \leq b_u. \end{aligned}$$

Dabei ist

$$\begin{aligned} H &= \begin{bmatrix} B_{pp} & B_{ps}^1 & B_{pu}^1 & \cdots & B_{ps}^N & B_{pu}^N \\ (B_{ps}^1)^\top & B_{ss}^1 & B_{su}^1 & & & \\ (B_{pu}^1)^\top & (B_{su}^1)^\top & B_{uu}^1 & & & \\ \vdots & & & \ddots & & \\ (B_{ps}^N)^\top & & & & B_{ss}^N & B_{su}^N \\ (B_{pu}^N)^\top & & & & (B_{su}^N)^\top & B_{uu}^N \end{bmatrix}, f = \begin{bmatrix} f_p \\ f_s^1 \\ f_u^1 \\ \vdots \\ f_s^N \\ f_u^N \end{bmatrix}, \\ S &= \begin{bmatrix} X_p^0 & X_s^0 & X_u^0 & -I & & \\ X_p^1 & & X_s^1 & X_u^1 & -I & \\ \vdots & & & \ddots & \ddots & \\ X_p^{N-1} & & & X_s^{N-1} & X_u^{N-1} & -I \end{bmatrix}, s = \begin{bmatrix} X_c^0 \\ X_c^1 \\ \vdots \\ X_c^{N-1} \end{bmatrix}, \\ Q &= \begin{bmatrix} R_p^0 & R_s^0 & R_u^0 & & & \\ R_p^1 & & R_s^1 & R_u^1 & & \\ \vdots & & & & \ddots & \\ R_p^N & & & & & R_s^N & R_u^N \\ C_p & C_s^0 & C_u^0 & C_s^1 & C_u^1 & \cdots & C_s^N & C_u^N \end{bmatrix}, q = \begin{bmatrix} R_c^0 \\ R_c^1 \\ \vdots \\ R_c^N \\ C_c \end{bmatrix}, \\ u &= \begin{bmatrix} p_l \\ s_l^0 \\ u_l^0 \\ \vdots \\ s_l^N \\ u_l^N \end{bmatrix}, l = \begin{bmatrix} p_u \\ s_u^0 \\ u_u^0 \\ \vdots \\ s_u^N \\ u_u^N \end{bmatrix}. \end{aligned}$$

2.2. Klassischer Ansatz. Jede Blockzeile der Matrix S hat vollen Zeilenrang. Also liefert die Bedingung $Sx = s$ eine eindeutige Abhängigkeit der Variablen. Condensing Algorithmen nutzen die Struktur der Matrix S aus um die Anzahl der Variablen über diese Abhängigkeit zu eliminieren.

Bei dem klassischen Ansatz wird $Sx = s$ so umsortiert, dass

$$\begin{bmatrix} -I & & & \\ X_s^0 & \ddots & & \\ & \ddots & -I & \\ & & X_s^{N-1} & -I \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_{N-1} \\ s_N \end{bmatrix} = s - \begin{bmatrix} X_p^0 & X_s^0 & X_u^0 & & \\ X_p^1 & & X_u^1 & & \\ \vdots & & & \ddots & \\ X_p^{N-1} & & & & X_u^{N-1} \end{bmatrix} \begin{bmatrix} p \\ s_0 \\ u_0 \\ \vdots \\ u_{N-1} \end{bmatrix}$$

gilt. Dieses Gleichungssystem ist in Abhängigkeit von $(p, s_0, u_0, \dots, u_{N-1})$ eindeutig lösbar.

Der Nachteil dieses Ansatzes wird deutlich, wenn die Matrizen X_s^i für $i = 0, \dots, N-1$ schlecht konditioniert sind. Da bei der Lösung des Gleichungssystem die Matrizen X_s^i auf multipliziert werden, treten bei Lösung numerische Schwierigkeiten auf.

2.3. Iterative QR-Faktorisierung und Variablentransformation. Um dies auf eine numerisch stabile Weise zu machen, nehmen wir eine orthogonale Variablentransformation vor. Dazu betrachten die einzelnen Matchingblöcke und berechnen eine entsprechende QR-Zerlegung.

Sei Q_0 und R_0 die QR-Zerlegung des Blockes $[X_s^0 X_u^0 - I]$, wobei Q_0 eine orthonormale Matrix der Dimension $n_u + 2n_s$ ist und R_0 eine invertierbare untere Dreiecksmatrix mit der Dimension n_s ist.

Es gilt also :

$$\begin{bmatrix} X_s^0 & X_u^0 & -I \\ 0 & 0 & X_s^1 \end{bmatrix} = \begin{bmatrix} R_0 & 0 & 0 \\ \tilde{X}_s^0 & \tilde{X}_u^0 & \hat{X}_s^1 \end{bmatrix} Q_0$$

Diese Faktorisierung wiederholen wir iterativ für alle Blockzeilen. Für $i = 2, \dots, N-1$ sind die Matrizen R_i und Q_i rekursiv so gewählt, dass

$$\begin{bmatrix} \hat{X}_s^i & X_u^i & -I \\ 0 & 0 & X_s^{i+1} \end{bmatrix} = \begin{bmatrix} R_i & 0 & 0 \\ \tilde{X}_s^i & \tilde{X}_u^i & \hat{X}_s^{i+1} \end{bmatrix} Q_i$$

gilt. Sei \tilde{Q}_i die orthonormale Matrix, die die entsprechende Spaltenoperation auf ganz S repräsentiert. Die Matrix hat also die Struktur

$$\tilde{Q}_i = \left[\begin{array}{c|c|c} I_{n_p+(i-1)(n_s+n_u)} & & \\ \hline & Q_i & \\ \hline & & I_{(N-i-2)(n_u+n_s)-n_s} \end{array} \right].$$

Die orthonormale Variablentransformation ist also durch die Matrizen \tilde{Q}_i gegeben. Somit können wir gilt für S :

$$\begin{aligned} S &= \begin{bmatrix} X_p^0 & R_0 & 0 & 0 & & & \\ X_p^1 & \tilde{X}_s^0 & \tilde{X}_u^0 & \hat{X}_s^1 & X_u^1 & -I & \\ \vdots & & & & \ddots & \ddots & \ddots \\ X_p^{N-1} & & & & X_s^{N-1} & X_u^{N-1} & -I \end{bmatrix} \tilde{Q}_0 \\ &= \begin{bmatrix} X_p^0 & R_0 & & & & & \\ X_p^1 & \tilde{X}_s^0 & \tilde{X}_u^0 & R_1 & & & \\ X_p^2 & & \tilde{X}_s^1 & \tilde{X}_u^1 & \hat{X}_s^2 & X_u^2 & -I \\ \vdots & & & & \ddots & \ddots & \ddots \\ X_p^{N-1} & & & & X_s^{N-1} & X_u^{N-1} & -I \end{bmatrix} \tilde{Q}_1 \tilde{Q}_0 \\ &= \underbrace{\begin{bmatrix} X_p^0 & R_0 & & & & & \\ X_p^1 & \tilde{X}_s^0 & \tilde{X}_u^0 & R_1 & & & \\ X_p^2 & & \tilde{X}_s^1 & \tilde{X}_u^1 & R_2 & & \\ \vdots & & & \ddots & \ddots & \ddots & \\ x_p^{N-1} & & & R_{N-1} & 0 & 0 \end{bmatrix}}_{:=\tilde{S}} \underbrace{\tilde{Q}_{N-1} \dots \tilde{Q}_1 \tilde{Q}_0}_{:=Q}. \end{aligned}$$

Anschließend sortieren wir die Spalten von \tilde{S} so um dass $\hat{S} = [\hat{S}_1 | \hat{S}_2]$ gilt, mit

$$\hat{S}_1 = \begin{bmatrix} R_0 & & & & \\ \tilde{X}_s^0 & R_1 & & & \\ & \tilde{X}_s^1 & \ddots & & \\ & & \ddots & R_{N-2} & \\ & & & \tilde{X}_s^{N-2} & R_{N-1} \end{bmatrix}, \hat{S}_2 = \begin{bmatrix} X_p^0 & \tilde{X}_u^0 & & & \\ X_p^1 & & \tilde{X}_u^1 & & \\ \vdots & & & \ddots & \\ X_p^{N-2} & & & & \tilde{X}_u^{N-2} \\ X_p^{N-1} & & & & \tilde{X}_u^{N-1} \end{bmatrix}.$$

Dabei ist \hat{S}_1 eine invertierbare Diagonalmatrix. Algebraisch wird die Umsortierung der Spalten durch die Multiplikation von rechts mit einer Permutationsmatrix P realisiert. Es gilt also

$$S = \tilde{S}Q = \hat{S}P^\top Q = [\hat{S}_1 | \hat{S}_2] P^\top Q.$$

Wir definieren dann die orthonormale Transformationsmatrix $T := P^\top Q$.

Wir benutzen diese Matrix um das QP zu transformieren. Dafür setzen wir $u = Tx$. Dann ist das QP äquivalent zu

$$(2.2) \quad \begin{aligned} \max_u \quad & \frac{1}{2} u^\top \hat{H} u + \hat{f}^\top u \\ \text{s.t.} \quad & [\hat{S}_1 | \hat{S}_2] u = s \\ & \hat{Q} u = q \\ & b_l \leq T^\top u \leq b_u \end{aligned}$$

mit $\hat{H} = THT^\top$, $\hat{f} = Tf$ und $\hat{Q} = QT^\top$.

2.4. Condensing. Um die Anzahl der Variablen zu reduzieren nutzen wir aus, dass die Matrix \hat{S}_1 invertierbar ist und diagonalstruktur hat. Sei $u = [v^\top w^\top]^\top$, so dass

$$[\hat{S}_1 | \hat{S}_2] u = \hat{S}_1 v + \hat{S}_2 w$$

gilt. Da \hat{S}_1 invertierbar ist folgt daraus, dass

$$v = \underbrace{\hat{S}_1^{-1} s}_{:=s'} - \underbrace{\hat{S}_1^{-1} \hat{S}_2}_{:=S'} w$$

gilt. Substituieren wir v in dem QP 2.2, so erhalten wir folgendes QP:

$$(2.3) \quad \begin{aligned} w^* = \max_w \quad & \frac{1}{2} w^\top (S'^\top H_{11} S' - H_{12}^\top S' - S'^\top H_{12} + H_{22}) w \\ & + (H_{12}^\top s' - S'^\top H_{11} s' - S'^\top f_1 + f_2)^\top w \\ \text{s.t.} \quad & (Q_2 - Q_1 S') w = q - Q_1 s' \\ & b_l - T^\top \begin{pmatrix} s' \\ 0 \end{pmatrix} \leq T^\top \begin{pmatrix} -S' \\ I \end{pmatrix} w \leq b_u - T^\top \begin{pmatrix} s' \\ 0 \end{pmatrix}, \end{aligned}$$

wobei $\hat{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^\top & H_{22} \end{bmatrix}$, $\hat{f} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$ und $\hat{Q} = [Q_1 \ Q_2]$. Es gilt $w \in \mathbb{R}^{n_p + n_s + Nn_u}$. Das QP 2.3 hat also eine deutlich kleinere Dimension als 2.1. Dabei haben wir nur äquivalente Umformungen vorgenommen. Für jeden kritischen Punkt w von dem QP 2.3 erhalten wir also einen kritischen Punkt

$$x = T^\top \begin{bmatrix} s' - S' \\ I \end{bmatrix} w$$

von dem ursprünglichen QP 2.1. Mit der Annahme, dass die Matrix H positiv definit ist und somit die Lösung der QPs eindeutig ist, gilt:

$$x^* = T^\top \begin{bmatrix} s' - S' \\ I \end{bmatrix} w^*.$$

IWR UNI HEIDELBERG

E-mail address: `robert.scholz@iwr.uni-heidelberg.de`