Optimization of Power Transformations in Global Optimization

Andreas Lundell and Tapio Westerlund
Process Design Laboratory, Åbo Akademi University,
Biskopsgatan 8, FIN-20500 Åbo, Finland

A method for minimizing the number of power transformations needed or variables involved for the convexification of signomial functions in global optimization problems is presented in this article. By utilizing this method, the complexity of the transformed and convexified problem can be reduced and, hence, be solved more efficiently. The method is based on the fact that signomial terms can always be convexified by applying power transformations to the individual variables included in the terms. Furthermore, by properly selecting these power transformations and approximating the inverse transformations with piecewise linear functions, the feasible region of the problem can be overestimated, making it possible to solve optimization problems of the mentioned form to global optimality as a sequence of convex mixed integer programming problems.

1. Convexification and Underestimation of Signomial Terms

A *signomial function* is defined as the sum of signomial terms, where each term consists of power functions, *i.e.*

$$\sigma(\mathbf{z}) = \sum_{j=1}^{J} c_j \prod_{i=1}^{N} z_i^{p_{ji}}, \quad \mathbf{z} = (z_1, z_2, \ldots, z_N)\,. \tag{1}$$

It is clear that a signomial function is convex if the terms are convex. There are many different ways to convexify a non-convex signomial function, for example, by the use of exponential, inverse or power transformations. The latter are utilized in this article, and their validity is based on results from Maranas and Floudas (1995) stating that:

A) A *positive signomial term* is convex if:
   (i) all the powers are negative, or
   (ii) one power is positive and the rest negative, and the sum of the powers is greater than or equal to one.
B) A *negative signomial term* is convex if all powers are positive and the sum of the powers is smaller than or equal to one.

From these statements, it can be deduced that an arbitrary signomial term can always be convexified by using a power transformation $z = Z^Q$ and its inverse transformation $Z = z^{1/Q}$, where the variables Q fulfill certain criteria. Applying this transformation to the variables in the term does indeed convexify the signomial term, but only by moving the non-convexities from the signomial terms to the constraints introduced by the inverse transformations. However, by approximating the non-linear inverse transformations with piecewise linear functions, the whole problem can be convexified, on the condition that the approximation of each transformed signomial term underestimates the original term. This is accomplished by introducing additional requirements on the transformations. The power transformations and the underestimations mentioned have been studied previously, for example, in Pörn (2000), Björk (2002) and Westerlund (2005).

Since the convexity requirements are different for negative and positive signomial terms, different requirements on Q are needed. For negative terms ($c < 0$), the power Q must be positive for variables with positive powers and negative for variables with negative powers. Also, if the transformed terms are to underestimate the original term, certain conditions on the variables Q must be satisfied. For negative terms, these conditions coincide with the convexification requirements, so the statements

$$\begin{cases} Q_i > 0, & \forall i : p_i > 0, \\ Q_i < 0, & \forall i : p_i < 0, \end{cases} \qquad \text{and} \qquad \sum_{i=1}^{N} p_i Q_i \leq 1 \tag{2}$$

must be fulfilled, if the power transformations are to give valid convex underestimates. For a non-convex positive signomial term, the convexity requirements on the powers Q are, that at most, one positive power may remain positive after the transformation, *i.e.* there may exist an index k ($1 \leq k \leq N$), so that the product $p_k Q_k$ is positive, while $p_i Q_i$ is negative for all other indices $i \neq k$. The variables with negative powers do not need to be transformed at all, so Q is defined to be one in this case. By combining these requirements with the requirements needed for underestimation, the conditions

$$\begin{cases} Q_i \geq 1, & \forall i : p_i > 0 \quad \wedge \quad i = k, \\ Q_i < 0, & \forall i : p_i > 0 \quad \wedge \quad i \neq k, \\ Q_i = 1, & \forall i : p_i < 0, \end{cases} \qquad \text{and} \qquad \sum_{i=1}^{N} p_i Q_i \geq 1 \tag{3}$$

are received for positive signomial terms (the sum is only required if an index k exists). Using power transformations of the mentioned form, signomial terms can always be convexified and underestimated whenever the inverse transformations are approximated with piecewise linear functions.

2. Optimization of the Power Transformations

The power transformation method used in the above convexification procedure, can be optimized by making sure that Q is equal to one, *i.e.* no transformation takes place, whenever allowed by the convexification and underestimation requirements. Presented here is a method for the convexification of a problem with J_T non-convex signomial terms, see Lundell (2007). From now on, the indices j correspond, not only to non-convex signomial terms in a single signomial function, but to all non-convex signomial terms in inequality constraints less than or equal to zero in the problem.

Since $z_i^{(1+b_{ji}(Q_{ji}-1))p_{ji}}$ simplifies to $z_i^{p_{ji}}$ if b_{ji} is equal to zero and to $z_i^{Q_{ji}p_{ji}}$ if b_{ji} is equal to one, the introduction of a binary variable b_{ji}, being zero if the variable z_i in the j-th term is not transformed, and one otherwise, makes it possible to write the transformed signomial terms as $c_j \prod_{i=1}^{N} z_i^{(1+b_{ji}(Q_{ji}-1))p_{ji}}$. A MILP problem can then be formulated, with the objective being to minimize the number of transformations required for the convexification and underestimation of the J_T signomial terms in the problem, *i.e.*

$$\min\left\{\sum_{j=1}^{J_T}\sum_{i=1}^{N} b_{ji}\right\}, \tag{4}$$

under the conditions that the powers Q satisfy the requirements (2) or (3), depending on whether the term in question is negative or positive. The solution to this problem will indicate, not only the number of transformations required, but also which variables need to be transformed, as well as, the power transformations that can be used.

However, when approximating each inverse power transformation with a piecewise linear function, the same binary variables can be used in all piecewise linear approximations of the inverse transformation of the same original variable z_i, even if the transformations used are not the same. Therefore, it could be more beneficial to minimize the total number of original variables involved in the transformations rather than the total number of transformations, hereby also minimizing the number of binaries needed in the piecewise linear functions. By introducing a new binary variable B_i, equal to one if the i-th variable is transformed by a power transformation in any of the terms where it is found, and zero otherwise, it is possible to minimize the number of original variables involved in any of the transformations. Since it is still important to keep the total number of transformations down, the previous objective function (4) is included, but multiplied with a small positive value δ, to give the minimization of the number of transformations less weight than that of the transformed variables. Furthermore, to promote numerically more stable powers of Q, an additional penalty term (multiplied with δ^2) consisting of the sum of the deviations Δ of the powers Q from +1 if Q is positive (when the binary $\beta=1$) and -1 if Q is negative (when the binary $\beta=0$). Hence, the new objective, as well as, the conditions on the binaries B and the deviations Δ can be written as

$$\min\left\{\sum_{i=1}^{N} B_i + \delta\sum_{j=1}^{J_T}\sum_{i=1}^{N} b_{ji} + \delta^2\sum_{j=1}^{J_T}\sum_{i=1}^{N}\Delta_{ji}\right\}, \quad \text{subject to} \tag{5}$$

$$\begin{cases} \sum_{j=1}^{J_T} b_{ji} \le J_T B_i, \\ Q_{ji} - \Delta_{ji} + M\beta_{ji} \le M+1, \\ -Q_{ji} - \Delta_{ji} + M\beta_{ji} \le M-1, \end{cases} \quad \text{and} \quad \begin{cases} Q_{ji} - \Delta_{ji} - M\beta_{ji} \le -1, \\ -Q_{ji} - \Delta_{ji} - M\beta_{ji} \le 1, \\ M(\beta_{ji}-1) \le Q_{ji} \le M\beta_{ji}. \end{cases} \tag{6}$$

The conditions guaranteeing correct power transformations also need to be added to the linear problem. These conditions, given below, are different depending on the sign of the signomial terms, *i.e.* whether c_j is positive or negative in the j-th signomial term.

2.1. Conditions for the negative signomial terms

For negative signomial terms ($c_j < 0$), it must be guaranteed that, when the power is positive ($p > 0$) and a transformation is necessary ($b = 1$), then Q must be between zero and one, and if a transformation is not needed ($b = 0$), then Q is equal to one. These conditions can be formulated as the linear inequalities (7), where the index i corresponds to the indices for the positive powers in the negative signomial j.

$$\begin{cases} Q_{ji} \geq 1 - b_{ji}, \\ Q_{ji} \leq 1 - \varepsilon \cdot b_{ji}, \\ Q_{ji} \geq \varepsilon, \end{cases} \Rightarrow \begin{cases} b_{ji} = 0: & Q_{ji} = 1, \\ b_{ji} = 1: & \varepsilon \leq Q_{ji} \leq 1 - \varepsilon. \end{cases} \tag{7}$$

A positive constant $\varepsilon = 1/M$, where M is a large number, has been used in the inequalities (7) to give practical bounds. When the power is negative ($p < 0$), a transformation is always necessary and Q must be negative, so $b_{ji} = 1$ and $-M \leq Q_{ji} \leq -\varepsilon$, where the indices i correspond to the variable with negative powers in the negative signomial term j. Furthermore, the sum of all the powers in the signomial term after transformation must be less than or equal to one, so the requirement $\sum_{i=1}^{N} p_i Q_i \leq 1$ must also be fulfilled.

2.2. Conditions for the positive signomial terms

For positive signomial terms, more freedom exists regarding how the transformations can be chosen, since a term can be convexified in two different ways: either all the variables have negative powers after the transformation or one variable has a positive power and the rest of the variables have negative powers. To be able to handle the variables with positive powers, a binary variable α_{ji} is introduced. This variable is equal to one if z_i has a positive power after the transformation in the j-th term, and equal to zero otherwise. Since at most one transformed variable per term can have a positive power, the requirement $\sum_{i=1}^{N} \alpha_{ji} \leq 1$ is necessary.

The variable Q should be larger or equal to one for a variable with the positive transformation ($\alpha = 1$), and be smaller than zero for the rest of the originally positive powers (with $\alpha = 0$). Therefore, the following conditions must be included:

$$\begin{cases} Q_{ji} \leq \alpha_{ji} M - \varepsilon(1 - \alpha_{ji}), \\ Q_{ji} \geq -M + \alpha_{ji}(M + 1), \end{cases} \Rightarrow \begin{cases} \alpha_{ji} = 0: & -M \leq Q_{ji} \leq -\varepsilon, \\ \alpha_{ji} = 1: & 1 \leq Q_{ji} \leq M. \end{cases} \tag{8}$$

Also, for variables with positive powers ($p > 0$), the binary b should be equal to zero when no transformation occurs, *i.e.* when $\alpha = 1$ and $Q = 1$, and equal to one otherwise. Using the same values on M and ε as in section 2.1, this can be formulated as:

$$\begin{cases} b_{ji} \geq 1 - \alpha_{ji}, \\ b_{ji} \geq \varepsilon(Q_{ji} - 1), \\ b_{ji} \leq (1 - \varepsilon) Q_{ji} + M(1 - \alpha_{ji}), \end{cases} \Rightarrow \begin{cases} \alpha_{ji} = 0 \wedge Q_{ji} < 0: & b_{ji} = 1, \\ \alpha_{ji} = 1 \wedge Q_{ji} = 1: & b_{ji} = 0,, \\ \alpha_{ji} = 1 \wedge Q_{ji} \geq \frac{1}{1-\varepsilon}: & b_{ji} = 1. \end{cases} \tag{9}$$

where i corresponds to the indices for the variables with the positive powers in the j-th signomial term. For the negative powers ($p < 0$) a transformation is not needed, so the following conditions are included in the linear problem: $b_{ji} = 0$, $Q_{ji} = 1$ and $\alpha_{ji} = 0$. Furthermore, the sum of all the powers in the signomial term after transformation should be greater or equal to one if there exists one variable with positive power, *i.e.* $\sum_{i=1}^{N} \alpha_{ji} = 1$, and smaller than zero otherwise. Hence, $\sum_{i=1}^{N} p_i Q_i - M \sum_{i=1}^{N} \alpha_{ji} \geq 1 - M$ must also be valid for a positive signomial term.

By formulating a mixed integer linear minimization problem with the objective function and linear constraints presented in this chapter, it is possible to determine not only which variables need to be transformed, but also the power transformations required for convexification and underestimation of all the signomial terms in the original problem.

3. Examples and Simulations of the Method

3.1. An example

The following geometric programming problem from Rijckaert and Martens (1978), including seven signomial terms and eight variables, is used to illustrate the procedure.

$$\text{MINIMIZE} \quad 2.0 \cdot z_1^{0.9} z_2^{-1.5} z_3^{-3} + \underline{5.0 \cdot z_4^{-0.3} z_5^{2.6}} + 4.7 \cdot z_6^{-1.8} z_7^{-0.5} z_8 ,$$

$$\text{s.t} \quad 7.2 \cdot z_1^{-3.8} z_2^{2.2} z_3^{4.3} + \underline{0.5 \cdot z_4^{-0.7} z_5^{-1.6}} + 0.2 \cdot z_6^{4.3} z_7^{-1.9} z_8^{8.5} \leq 1,$$

$$10.0 \cdot z_1^{2.3} z_2^{1.7} z_3^{4.5} \leq 1, \quad 0.6 \cdot z_4^{-2.1} z_5^{0.4} \leq 1, \quad \underline{6.2 \cdot z_6^{4.5} z_7^{-2.7} z_8^{-0.6}} \leq 1,$$

$$3.1 \cdot z_1^{1.6} z_2^{0.4} z_3^{-3.8} \leq 1, \quad 3.7 \cdot z_4^{5.4} z_5^{1.3} \leq 1, \quad \underline{0.3 \cdot z_6^{-1.1} z_7^{7.3} z_8^{-5.6}} \leq 1.$$

The underlined terms are convex, and can be disregarded. The method was applied to the rest, resulting in a total of 12 power transformations being needed to convexify and underestimate the terms, and that the variables z_4 and z_6 do not require any transformation at all. The resulting convexified signomial terms and transformations are:

Before transformation	After transformation	Power transformations
$2.0 \cdot z_1^{0.9} z_2^{-1.5} z_3^{-3}$	$2.0 \cdot Z_{1,1}^{-0.9} z_2^{-1.5} z_3^{-3}$	$z_1 = Z_{1,1}^{-1}$
$4.7 \cdot z_6^{-1.8} z_7^{-0.5} z_8$	$4.7 \cdot z_6^{-1.8} z_7^{-0.5} Z_{2,8}^{-1}$	$z_8 = Z_{2,8}^{-1}$
$7.2 \cdot z_1^{-3.8} z_2^{2.2} z_3^{4.3}$	$7.2 \cdot z_1^{-3.8} Z_{3,2}^{-2.2} Z_{3,3}^{7}$	$z_2 = Z_{3,2}^{-1}$, $z_3 = Z_{3,3}^{-1.6279}$
$0.2 \cdot z_6^{4.3} z_7^{-1.9} z_8^{8.5}$	$0.2 \cdot z_6^{4.3} z_7^{-1.9} Z_{4,8}^{-1.4}$	$z_8 = Z_{4,8}^{-0.16471}$
$10.0 \cdot z_1^{2.3} z_2^{1.7} z_3^{4.5}$	$10.0 \cdot Z_{5,1}^{-1.8} Z_{5,2}^{-1.7} z_3^{4.5}$	$z_1 = Z_{5,1}^{-0.78261}$, $z_2 = Z_{5,2}^{-1}$
$0.6 \cdot z_4^{-2.1} z_5^{0.4}$	$0.6 \cdot z_4^{-2.1} Z_{6,5}^{-0.4}$	$z_5 = Z_{6,5}^{-1}$
$3.1 \cdot z_1^{1.6} z_2^{0.4} z_3^{-3.8}$	$3.1 \cdot Z_{7,1}^{5.2} Z_{7,2}^{-0.4} z_3^{-3.8}$	$z_1 = Z_{7,1}^{3.25}$, $z_2 = Z_{7,2}^{-1}$
$3.7 \cdot z_4^{5.4} z_5^{1.3}$	$3.7 \cdot z_4^{5.4} Z_{8,5}^{-1.3}$	$z_5 = Z_{8,5}^{-1}$
$0.3 \cdot z_6^{-1.1} z_7^{7.3} z_8^{-5.6}$	$0.3 \cdot z_6^{-1.1} Z_{9,7}^{-7.3} z_8^{-5.6}$	$z_7 = Z_{9,7}^{-1}$

Approximating the inverse power transformations with piecewise linear functions, for example, using the SOS 2 representation discussed in Beale and Forrest (1976), the signomial terms will be underestimated.

3.2. Simulations of the method used on randomly generated signomial terms

To determine how well the method works, simulations on two different problem sizes were performed. Each simulation consisted of 50,000 groups of random signomial terms: one with 25 terms and 50 variables per group and one with 50 terms and 100 variables per group. The signomial terms (both negative and positive, as well as, convex and non-convex) were generated so that most of them were products of one, two or

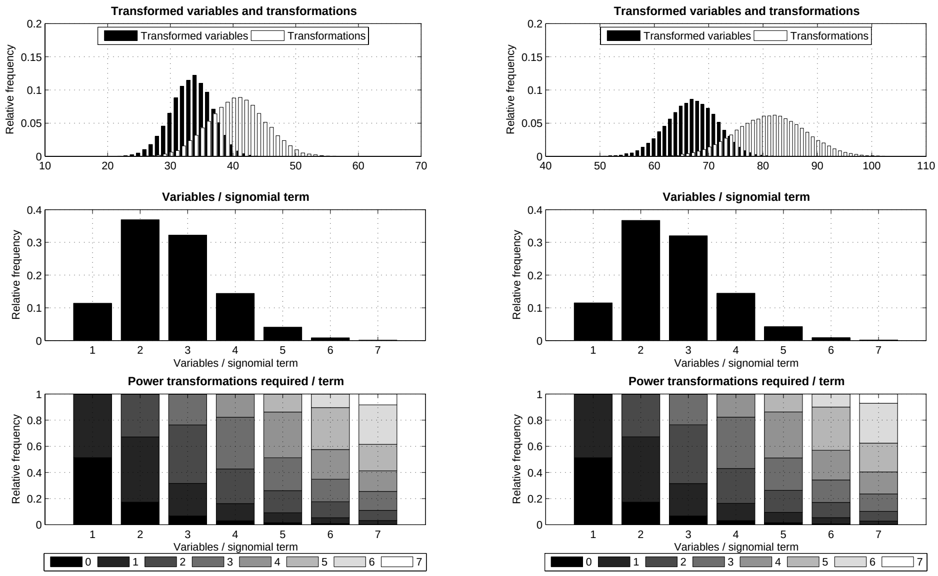

Figure 1: Simulations with 50,000 randomly generated groups of signomials with 25 terms and 50 variables (left) and 50 terms and a total of 100 variables (right).

three variables, and no term of more than seven variables. The powers were normally distributed with mean zero and variance four. The results are presented in Figure 1, from where it can be seen that often not all variables need to be transformed for the function to be convexified. The median number of transformed variables in the smaller simulation was 34, and in the larger simulation 67, so in both cases about one third of the original variables did not need to be transformed at all. The median number of power transformations required was 41 in the smaller case and 82 in the larger case.

4. References

Beale, E.M.L. and Forrest, J.J.H. (1976). Global Optimization Using Special Ordered Sets. Mathematical Programming 10.

Björk, K.-M. (2002). A Global Optimization Method with some Heat Exchanger Network Applications. PhD-thesis, Åbo Akademi University.

Rijckaert, M.J. and Martens, X.M. (1978). Comparison of Generalized Geometric Programming Algorithms. Journal of Optimization Theory and Applications 26.

Lundell, A. (2007). Optimization Techniques in Global Optimization. Master's thesis, Åbo Akademi University.

Maranas, C.D. and Floudas, C.A. (1995). Finding All Solutions of Nonlinearly Constrained Systems of Equations. Journal of Global Optimization 7.

Pörn, R. (2000). Mixed Integer Non-Linear Programming: Convexification Techniques and Algorithm Development. PhD-thesis, Åbo Akademi University.

Westerlund, T. (2005). Some Transformation Techniques in Global Optimization. Chapter in: Global Optimization: from Theory to Implementation. L. Liberti and N. Maculan (Eds.). Springer, pp 47-74, ISBN 0-387-28260-2.