CHEMICAL ENGINEERING TRANSACTIONS Volume 17, 2009
Editor Sauro Pierucci
ISBN 978-88-95608-01-3 ISSN 1974-9791
DOI: 10.3303/CET0917215

On the Relationship between Power and Exponential Transformations for Positive Signomial Functions

Andreas Lundell and Tapio Westerlund
Process Design & Systems Engineering, Åbo Akademi University
Biskopsgatan 8, FIN-20500 Turku, Finland

Global optimization of mixed integer nonlinear programming (MINLP) problems containing signomial terms is in many cases a difficult task, and many different approaches to solve these problems have been devised. In Westerlund (2005) a method where a relaxed convex relaxation of the original problem is obtained by approximating single-variable transformations with piecewise linear functions (PLFs) is described. What transformations are applicable depend on the sign of the signomial terms. For positive terms power transformations (PTs) (Lundell et al., 2007) and the exponential transformation (ET) (Lundell and Westerlund, 2008) can be used. In Lundell and Westerlund (2009) it was shown that the ET always results in a tighter underestimator than the negative power transformation (NPT). These results are in this paper extended to also include the positive power transformation (PPT); we will show that the PPT is generally better than the NPT, and also describe the relationship between the PPT and the ET.

Introduction

A special group of functions appearing in mathematical models of many processes is the signomial functions. A signomial function is defined as the sum of signomial terms, which in turn are products of power functions multiplied with a real constant, *i.e.*,

$$\sigma(\boldsymbol{x}) = \sum_j c_j \sum_i x_i^{p_{ji}}.$$

The definition of a signomial function provides a very general formulation incorporating many different types of functions; for example, bi- and trilinear terms, as well as all polynomials can be regarded as special cases of signomials. The global optimization method described in, *e.g.*, Westerlund (2005) and Lundell et al. (2007) can be applied to MINLP problems of the form

$$\begin{array}{lll}
\text{minimize} & f(\boldsymbol{x}), & \boldsymbol{x} = (x_1, x_2, \ldots, x_I), \\
\text{subject to} & \mathbf{A}\boldsymbol{x} = \boldsymbol{a}, & \mathbf{B}\boldsymbol{x} \le \boldsymbol{b}, \\
 & \boldsymbol{g}(\boldsymbol{x}) \le \mathbf{0}, & \boldsymbol{q}(\boldsymbol{x}) + \boldsymbol{\sigma}(\boldsymbol{x}) \le \mathbf{0},
\end{array}$$

Please cite this article as: Lundell A. and Westerlund T., (2009), On the relationship between power and exponential transformations for positive signomial functions, Chemical Engineering Transactions, 17, 1287-1292 DOI: 10.3303/CET0917215

where the functions f, g and q are convex, the functions σ are signomial and $x_i > 0\ \forall i$. When approximating the inverse transformations with PLFs, the whole MINLP problem can be written in a convex relaxed form overestimating the original problem, and by iteratively improving the approximations the global optimal solution can be found.

The transformation approach

A positive nonconvex signomial term can be convexified by applying the single-variable transformations $x_i = e^{X_i}$ (ET) or $x_i = {X_i}^{Q_i}$ (PT) to the variables having positive powers p_i. The power transformations can be of two different types, the negative power transformation (NPT) or the positive power transformation (PPT). To guarantee a transformation resulting in convex terms, in the NPT all powers Q_i must be negative and in the PPT all powers except one (Q_k) must be negative and the sum of the powers must be greater than or equal to one, *i.e.*,

$$\sum_{i:p_i > 0} p_i Q_i + \sum_{i:p_i < 0} p_i \geq 1.$$

In addition to introducing the transformations mentioned above, approximations in the form of PLFs of the inverse transformations $X_i = \ln x_i$ and $X_i = x_i^{1/Q_i}$ for the ET and PTs respectively, must be included in the MINLP problem. See Westerlund (2005) or Lundell and Westerlund (2009) for details.

Example 1. To illustrate how the transformation technique works, the ET, PPT and NPT are now applied to the signomial function $f(x, y) = xy$. Using the transformation $x = e^{X_E}$ in the ET, and the transformation $x = X_P^Q$ for the PTs, the transformed terms will then take the form:

ET: $xy \to e^{X_E} e^{Y_E}$, PPT: $xy \to X_P^{Q_{x,1}} Y_P^{Q_{y,1}}$, NPT: $xy \to X_P^{Q_{x,2}} Y_P^{Q_{y,2}}$.

In the PPT the transformation powers $Q_{x,1}$ and $Q_{y,1}$ must be chosen such that one is positive, the other negative, and their sum is greater than or equal to one, *e.g.*, $Q_{x,1} = 2$ and $Q_{y,1} = -1$. The conditions on the transformation powers in the NPT is that they are both negative, *e.g.*, $Q_{x,2} = Q_{y,2} = -1$. Furthermore, the inverse transformations $X_E = \ln x$, $Y_E = \ln y$ and $X_P = x^{1/Q_x}, Y_P = y^{1/Q_y}$ must be replaced by the PLFs $\hat{X}_E, \hat{Y}_E$ and $\hat{X}_P$, $\hat{Y}_P$. For the variable x in one step on the interval $[\underline{x}, \overline{x}]$ the PLFs are given as

$$\hat{X}_E(x) = \ln \underline{x} + \frac{\ln \overline{x} - \ln \underline{x}}{\overline{x} - \underline{x}}(x - \underline{x}) \quad \text{and} \quad \hat{X}_P(x) = \underline{x}^{1/Q} + \frac{\overline{x}^{1/Q} - \underline{x}^{1/Q}}{\overline{x} - \underline{x}}(x - \underline{x}).$$

Theoretical results

In this section we will show some new results regarding the relation between the PTs and the ET. This is a continuation of the work in Lundell and Westerlund (2009) where the following theorem was shown.

Theorem 1. The ET always gives a tighter convex underestimator when applied to an individual variable than the NPT regardless of the power Q used in the NPT.

The results in Theorem 1 can be extended to also include the PPT:

Theorem 2. A single-variable PPT always gives a tighter convex underestimation than both the ET and NPT.

Proof. This proof is similar to the proof for Theorem 1, which can be found in Lundell and Westerlund (2009). When the power function x^p is transformed by the transformations $x = X_P^Q, Q \geq 1$ (PPT) and $x = e^{X_E}$ (ET) and approximated with the PLFs $\hat{X}_P$ and $\hat{X}_E$ respectively, the claim that the PPT is a tighter underestimator than the ET is equivalent to

$$(\hat{X}_P^Q)^p \geq \left(e^{\hat{X}_E}\right)^p \overset{p>0}{\Longrightarrow} \hat{X}_P^Q \geq e^{\hat{X}_E}.$$

We now consider a one-step PLF only, so the previous inequality is equivalent to

$$\left(\underline{x}^{1/Q} + \frac{\overline{x}^{1/Q} - \underline{x}^{1/Q}}{\overline{x} - \underline{x}}\left(x - \underline{x}\right)\right)^Q \geq \exp\left(\ln \underline{x} + \frac{\ln \overline{x} - \ln \underline{x}}{\overline{x} - \underline{x}}\left(x - \underline{x}\right)\right).$$

By modifying this expression and introducing $z = \left(\frac{\overline{x}}{\underline{x}}\right)^{1/Q}$ and $\lambda = \frac{x - \underline{x}}{\overline{x} - \underline{x}}$, we receive the equivalent form $\lambda z \geq z^\lambda + \lambda - 1$, which can easily be proven to be true whenever $Q \geq 1$ and $\underline{x} \leq x \leq \overline{x}$ by searching for the minimum value of the function $f(z) = \lambda z - z^\lambda - \lambda + 1$.

Thus, the PPT gives a tighter convex underestimator than the ET when applied to individual variables. Furthermore, since we know from Theorem 1 that the ET gives better approximations than the NPT, it is obvious that the same is true for the PPT. It is also clear that the approximations are only equal in the endpoints $\underline{x}$ and $\overline{x}$. ■

Theorem 3. For the piecewise linear approximations $\hat{X}_P$ (PPT and NPT) and $\hat{X}_E$ (ET), the resulting convex underestimations of the PTs when the power Q goes to plus or minus infinity is equal to that of the ET, *i.e.*, $\lim_{Q \to \pm\infty} \hat{X}_P^Q = e^{\hat{X}_E}$.

Proof. The claim was proved for negative values of Q in Lundell and Westerlund (2009). The proof is, however, valid for positive values as well. ■

Although the PPT always results in a tighter convex underestimator than the ET when applied to an individual power function, as the following theorem shows, it is not generally true for the case when a PPT is applied to a signomial term of more than one variable.

Theorem 4. Neither the PPT nor the ET gives a tighter underestimator in the whole domain when applied to a nonconvex signomial term of more than one variable.

Proof. For a general nonconvex signomial term

$$s(\boldsymbol{x}) = \prod_{i \in I_+} x_i^{p_i} \cdot \prod_{i \in I_-} x_i^{p_i},$$

where $I_+ = \{i: p_i > 0\}$ and $I_- = \{i: p_i < 0\}$, only the variables $x_i: i \in I_+$ must be transformed. When using the ET all of the variables with positive powers are transformed using single-variable ETs and when using the PPT one variable (with arbitrary index $k \in I_+$) is transformed using a single-variable PPT and the rest (with indices $i \in I_+ \setminus \{k\}$) using single-variable NPTs. We know from Theorems 1 and 2 that

$$\hat{X}_{k,PT}^{Q_k p_k} \geq \left(\exp \hat{X}_{k,ET}\right)^{p_k} \qquad \text{and} \qquad \forall i \in I_+ \setminus \{k\}\text{:}\ \ \hat{X}_{i,PT}^{Q_i p_i} \leq \left(\exp \hat{X}_{i,ET}\right)^{p_i}.$$

when the transformation variables $X_{i,PT}$ and $X_{i,ET}$ have been replaced with the PLFs $\hat{X}_{i,PT}$ and $\hat{X}_{i,ET}$. Now, take two points $\boldsymbol{x}^* = (x_1^*, \dots, x_N^*)$ and $\boldsymbol{x}^\# = (x_1^\#, \dots, x_N^\#)$ such that:

$$x_i^*\text{:}\ \begin{cases} i = k\text{:} & \underline{x_i} < x_i^* < \overline{x}_i, \\ i \in I_+ \setminus \{k\}\text{:} & x_i^* = \underline{x_i} \vee x_i^* = \overline{x}_i\,, \\ i \in I_-\text{:} & \underline{x_i} \leq x_i^* \leq \overline{x}_i, \end{cases} \qquad x_i^\#\text{:}\ \begin{cases} i = k\text{:} & x_i^\# = \underline{x_i} \vee x_i^\# = \overline{x}_i, \\ i \in I_+ \setminus \{k\}\text{:} & \underline{x_i} < x_i^\# < \overline{x}_i, \\ i \in I_-\text{:} & \underline{x_i} \leq x_i^\# \leq \overline{x}_i. \end{cases}$$

Then, for the point $\boldsymbol{x}^*$ the following is true

$$\begin{aligned} \hat{s}_P(\boldsymbol{x}^*) &= \prod_{i \in I_+} \hat{X}_{i,PT}^{p_i Q_i} \cdot \prod_{i \in I_-} (x_i^*)^{p_i} = \hat{X}_{k,PT}^{p_k Q_k} \cdot \prod_{i \in I_+ \setminus \{k\}} \left(\exp \hat{X}_{i,ET}\right)^{p_i} \cdot \prod_{i \in I_-} (x_i^*)^{p_i} \\ &> \left(\exp \hat{X}_{k,ET}\right)^{p_k} \prod_{i: i \in I_+ \setminus \{k\}} \left(\exp \hat{X}_{i,ET}\right)^{p_i} \cdot \prod_{i \in I_-} (x_i^*)^{p_i} = \hat{s}_E(\boldsymbol{x}^*), \end{aligned}$$

since the underestimations resulting from the PPT always lie above those of the ET when $\underline{x_i} < x_i^* < \overline{x}_i$ and the NPT and the ET are equal at the endpoints, *i.e.*, when $x_i^* = \underline{x_i} \vee x_i^* = \overline{x}_i$. Therefore, the PPT gives a tighter underestimation than the ET in the chosen point . Correspondingly, for the point $\boldsymbol{x}^\#$ the following is true

$$\begin{aligned} \hat{s}_P(\boldsymbol{x}^\#) &= \prod_{i \in I_+} \hat{X}_{i,PT}^{p_i Q_i} \cdot \prod_{i \in I_-} \left(x_i^\#\right)^{p_i} = \left(\exp \hat{X}_{k,ET}\right)^{p_k} \cdot \prod_{i \in I_+ \setminus \{k\}} \hat{X}_{i,PT}^{p_i Q_i} \cdot \prod_{i \in I_-} \left(x_i^\#\right)^{p_i} \\ &< \left(\exp \hat{X}_{k,ET}\right)^{p_k} \cdot \prod_{i: i \in I_+ \setminus \{k\}} \left(\exp \hat{X}_{i,ET}\right)^{p_i} \cdot \prod_{i \in I_-} \left(x_i^\#\right)^{p_i} = \hat{s}_E(\boldsymbol{x}^\#). \end{aligned}$$

Thus, since both $\hat{s}_P(\boldsymbol{x}^*) > \hat{s}_E(\boldsymbol{x}^*)$ and $\hat{s}_P(\boldsymbol{x}^\#) < \hat{s}_E(\boldsymbol{x}^\#)$ are true, neither the PPT nor the ET gives tighter underestimations in the whole domain. Furthermore, since the underestimators are continuous, there exist parts of the domain where they are equal. ■

Corollary 1. The region where the PPT and the ET are equal consists of the part of the domain where the following expression is true

$$\prod_{i\in I_+} \hat{X}_{i,PT}^{p_i Q_i} = \prod_{i\in I_+} e^{p_i \hat{X}_{i,ET}}, \qquad I_+ = \{i: p_i > 0\}.$$

Illustrations and examples

Some illustrations of the convex underestimators arising from different transformations are now provided.

Example 2. The nonconvex function $f(x) = -8x + 0.05x^3 + 25x^{0.5},\ 1 \leq x \leq 7,$ consists of two convex terms as well as the nonconvex signomial term $25x^{0.5}$. Transforming the nonconvex term using the ET and the PTs and replacing the inverse transformation with PLFs, gives the convexified and underestimated functions

$$f(x, \hat{X}_E) = -8x + 0.05x^3 + 25e^{0.5\hat{X}_E} \text{ and } f(x, \hat{X}_P) = -8x + 0.05x^3 + 25\hat{X}_P^{0.5Q}.$$

The underestimators resulting from the ET and the PTs (with different values on the transformation power Q) are plotted in Figure 1. From the figure, it can be seen that the PPT and NPT gets closer to the ET when the power Q gets larger or smaller respectively, as stated in Theorem 3.

Example 3. Illustrations of the signomial function $f(x, y) = xy$ from Example 1 transformed using the ET and PPT ($Q_x = 2, Q_y = -1$) are provided in Figure 2. In Figure 3, the region where the PPT is a tighter underestimator than the ET is shown.

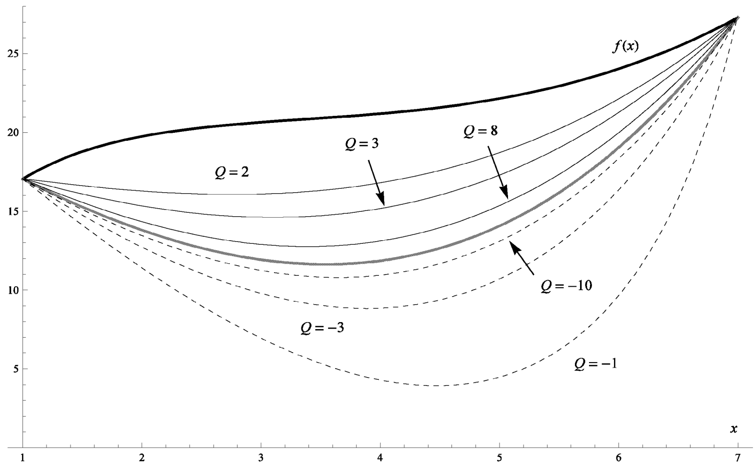

Figure 1 The function f(x) from Example 2 (thick, black) underestimated by the ET (thick, gray) and the PPT (black), as well as the NPT (dashed) for different values of the transformation power Q.

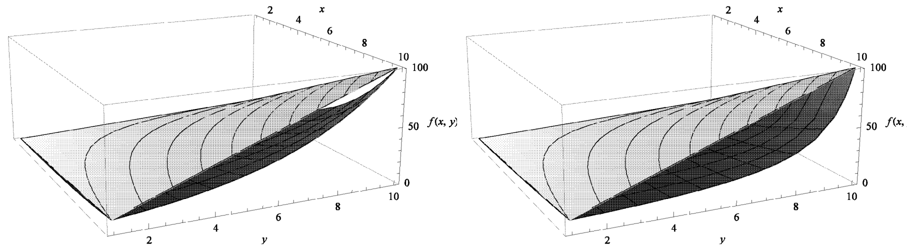

Figure 2. The convex underestimators resulting from the ET (left) and PPT (right) for the function f(x,y) in Example 3.

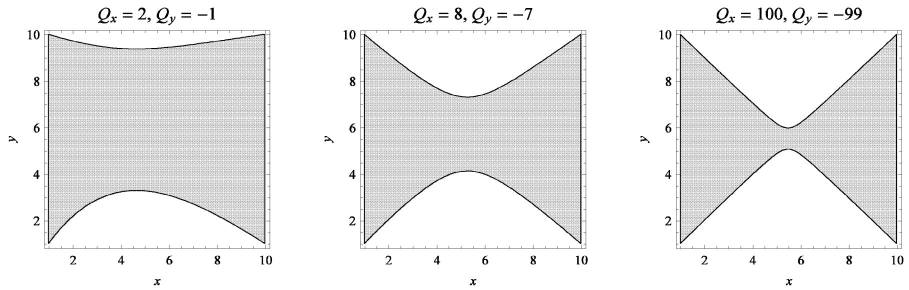

Figure 3. The shaded area indicates where the ET gives a better approximation than the PPT (the values of the transformation powers are indicated in the figure). The area has been obtained using the expression in Corollary 1.

Conclusions

In this paper, it was shown that the PPT gives a tighter convex underestimator than both the NPT and the ET when applied to an individual variable. Generally, however, the PPT applied to signomial terms of more than one variable, only gives a tighter underestimator than the ET in parts of the domain. An implicit expression showing in which parts of the domain the ET is superior is easy to deduce (Corollary 1), however, at this stage it is still not known whether explicit conditions exists.

References

Lundell A., Westerlund J. and Westerlund T., 2007 (accepted), Some transformation techniques with applications in global optimization, Journal of Global Optimization.

Lundell A. and Westerlund T., 2008, Exponential and power transformations for convexifying signomial terms in MINLP problems, Proceedings for the 27th IASTED MIC Conference.

Lundell A. and Westerlund T., 2009 (accepted), Convex underestimation strategies for signomial functions, Optimization Methods & Software.

Westerlund T., 2005, Global optimization: From theory to implementation: Some transformation techniques in global optimization, Eds. Liberti L. and Maculan N., Springer, New York.