

Datengenerierung für den Einsatz von Surrogatmodellen zur Approximation von thermodynamischen Gleichgewichten

Joschka Winz¹; Sebastian Engell¹

¹TU Dortmund, Fakultät Bio- und Chemieingenieurwesen, Lehrstuhl für Systemdynamik und Prozessführung, Deutschland

Abstract

In vielen Modellen chemischer Prozesse ist die akkurate Beschreibung der Thermodynamik von zentraler Bedeutung. Moderne thermodynamische Modelle, wie die Perturbed-Chain Statistical Associating Fluid Theory (PC SAFT), ermöglichen die akkurate Beschreibung, sind allerdings berechnungsaufwendig, was ein Hindernis zum Beispiel in der Optimierung von Prozessmodellen bedeutet.

Um diese Herausforderung zu überwinden, können Surrogatmodelle benutzt werden, welche die thermodynamischen Gleichgewichtsberechnungen approximieren und schnell auswertbar sind.

Der Einsatz von Surrogatmodellen führt zu einem Approximationsfehler. Da das Auswerten der zu approximierenden Funktion zeitaufwendig ist, ist es notwendig, mit möglichst wenig Trainingsdaten auszukommen. Verschiedene Ansätze stehen zur Verfügung.

Für das Ziel ein global genaues Surrogatmodell zu erzeugen, kann adaptives Sampling eingesetzt werden, bei dem die Vorhersage der Surrogatmodelle benutzt wird, um die Datengenerierung dort zu konzentrieren, wo der Prädiktionsfehler hoch ist. Ein weiterer Ansatz ist die Datenaugmentation, bei der aus einem bestehenden Set an Trainingsdaten weitere Daten generiert werden. Zum Beispiel kann im Falle von Gleichgewichtsdaten ausgenutzt werden, dass alle Feedzusammensetzungen einer Konode dieselben Phasenzusammensetzungen ergeben.

Im Fall der Prozessoptimierung ist es von zentraler Bedeutung, dass das Surrogatmodell am Optimum genau ist. Daher werden alternative Samplingmethoden benutzt um Trainingsdaten in dem Bereich, in dem das Prozessoptimum liegt, zu konzentrieren.

Diese Arbeit zeigt die möglichen Einsparungen an Funktionsaufrufen eines thermodynamischen Modells durch adaptives Sampling und Datenaugmentation in verschiedenen Zielsetzungen. Weiterhin werden Schwierigkeiten und Herausforderungen bei der Anwendung dieser Methoden präsentiert.