



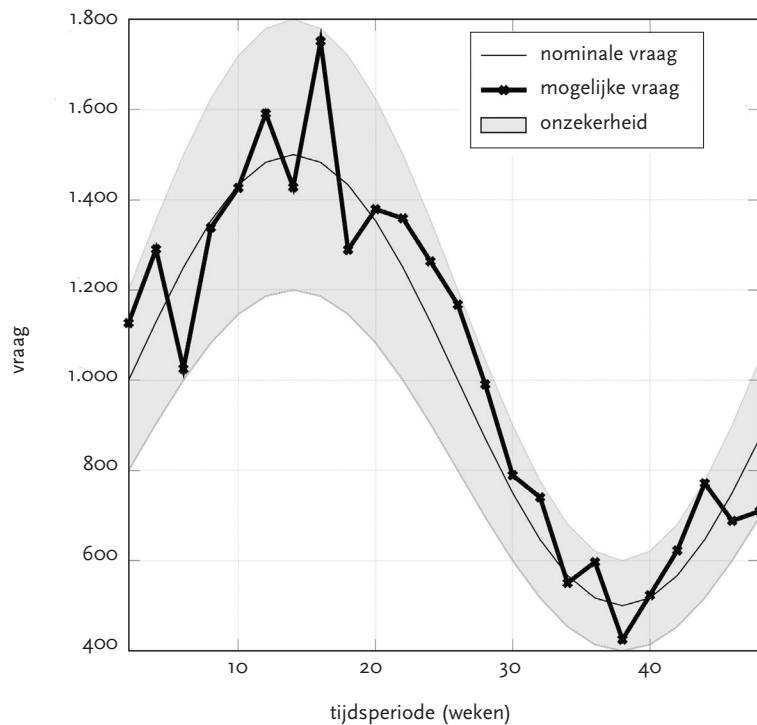
Wat te doen als gerapporteerde data inexact zijn?

FRANS DE RUITER

Bij productie-voorraadproblemen is de vraag naar de producten vaak onzeker. Dit maakt het opstellen van een productieplan dat aan alle vraag voldoet lastig, zeker als het plan moet voldoen aan restricties op productie- en voorraadniveaus. Het is door die restricties niet mogelijk om de productie voor de hele planningshorizon al van tevoren vast te stellen. Dit hoeft ook niet, omdat we elke periode kunnen bepalen hoeveel we gaan produceren. Dit probleem wordt daarom een *multi-stage probleem* genoemd. In dit multi-stage probleem is op tijdstip t bekend wat de vraag in alle voorgaande perioden 1,2 tot en met $t-1$ is geweest. Multi-stage technieken gebruiken deze data om de productie voor periode t te bepalen. Al deze technieken nemen echter aan dat de verkregen data exact zijn. In de praktijk blijkt dat deze data vaak inexact zijn: de werkelijke vraag verschilt van de gerapporteerde waarde. De inexacte aard van de gerapporteerde data kan grote gevolgen hebben voor het productieplan. In De Ruiter et al. (2014) is een methode ontwikkeld om dergelijke multi-stage voorraadproblemen met inexact verkregen data op te lossen.

Een toepassing: het productie-voorraad probleem

Ter illustratie van onze methode gebruiken we een bekend voorraadmodel uit het artikel van Ben-Tal, El Ghaoui en Nemirovski (2004). Het model bestaat uit één opslagfaciliteit en drie fabrieken die het product produceren. Om de week neemt het management een beslissing over de productie. In totaal moet de productie gepland worden voor 48 weken. De vraag naar de producten is seizoensgebonden, met een maximum vraag in de winter. Voor ons voorbeeld hebben we verschillende onzekerheidsniveaus getest van 2,5%, 5%, 10% en 20%. Een onzekerheidsniveau van 20% wil zeggen dat de werkelijke vraag tot 20% meer of minder kan zijn dan de door de manager verwachte vraag, de zogenaamde nominale vraag. Dit onzekerheidsniveau wordt opgesteld in samenspraak met het management en indien mogelijk aan de hand van historische data. Het productieplan moet zo opgesteld worden dat men aan alle vraag in het onzekerheidsgebied kan voldoen. Zie figuur 1 voor een grafische weergave van dit onzekerheidsgebied.



Figuur 1. Onzekerheid in de vraag. De grijs gekleurde strip geeft de onzekerheid van 20% weer.

Hierbij moeten we ook rekening houden met productiecosten, productiecapaciteit per fabriek, een minimum voorraadniveau en de opslagcapaciteit.

Het robuuste optimalisatiemodel

Alle restricties in het optimalisatieprobleem voor het productie-voorraadprobleem kunnen als lineaire ongelijkheden worden geschreven. Helaas weten we de vraag in ons model niet precies en wijkt deze af van de nominale waarde. Meestal is de informatie over de data erg miniem en kunnen we daarom niet vertrouwen op stochastische technieken. Daarom baseren we ons hier op robuuste optimalisatietechnieken waarbij het productieplan moet voldoen aan de restricties voor alle mogelijke vraagpatronen in het onzekerheidsgebied. Het eerste artikel dat dit idee ter sprake bracht is geschreven door Soyster (1978), maar intensief onderzoek naar robuuste optimalisatietechnieken kwam pas eind jaren 90 op gang door een reeks artikelen van de grondleggers van het onderzoeksveld: Aharon Ben-Tal, Laurent El Ghaoui en Arkadi Nemirovski. Diezelfde onderzoekers hebben recentelijk een boek geschreven (Ben-Tal et al., 2009) waarin een grote verscheidenheid aan robuuste optimalisatietechnieken wordt uitgelegd.

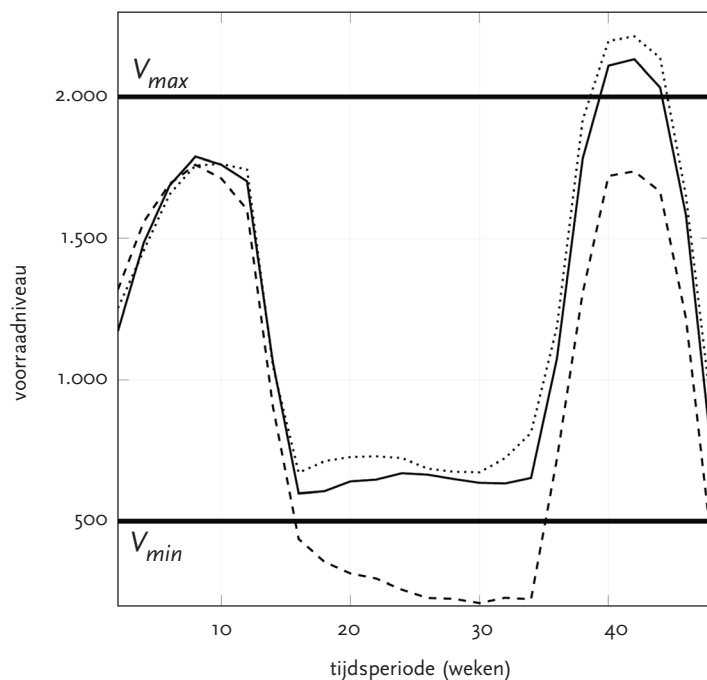
Robuust dynamische oplossingen

In ons productie-voorraadprobleem hoeven sommige beslissingen pas genomen te worden als een deel van de

onzekere data bekend is. Voor deze *wait-and-see* beslissingen in periode t kunnen we onze beslissingen baseren op vraagcijfers die beschikbaar zijn in de perioden 1 tot en met $t-1$. De *wait-and-see* variabelen in periode t zijn dan een wiskundige functie van de beschikbare data op dat moment. Dit model heet het *Adjustable Robust Optimization* (ARO) model (Ben-Tal et al., 2014). In dit model bedoelen we met *adjustable* dat sommige variabelen zich aanpassen aan de geobserveerde data. Vaak maken deze *wait-and-see* variabelen het probleem vele malen complexer, maar voor sommige gevallen kunnen we deze problemen nog efficiënt oplossen. Een van die gevallen is wanneer er geen onzekerheid is in de coëfficiënten van de *wait-and-see* variabelen en de mogelijke wiskundige functies voor deze variabelen lineair zijn. De coëfficiënten van deze lineaire functies zijn dan nieuwe optimalisatievariabelen die bepaald worden door ons model. De waarde voor de *wait-and-see* variabele kan dan bepaald worden zodra de data beschikbaar zijn. Met ARO-technieken kan dit model herschreven worden tot een efficiënt oplosbaar probleem zoals beschreven in Ben-Tal et al. (2014). In dit geval is het resulterende model een lineair programmeringsprobleem dat efficiënt op te lossen is met simplex of inwendige-punt methoden.

Inexacte data in de praktijk.

Inexacte data zijn nog steeds een groot probleem in de praktijk. Zo heeft volgens recent onderzoek door Haug et al. (2011) nog niet de helft van de bedrijven vertrouwen in



Figuur 2. Het negeren van de inexacte aard van data heeft gevolgen voor het voorraadniveau.

hun data. Nog belangrijker, 75% procent van de bedrijven heeft kosten kunnen identificeren die terug te leiden zijn op afwijkingen in de data. Voor voorraadmodellen zijn er zelfs studies, zoals DeHoratius en Raman (2008), die zeggen dat 65% van de gebruikte data inaccuraat zijn. Samenvattend kunnen we vaststellen dat verkregen data vaak inexact zijn.

Gevolgen van inexact verkregen data

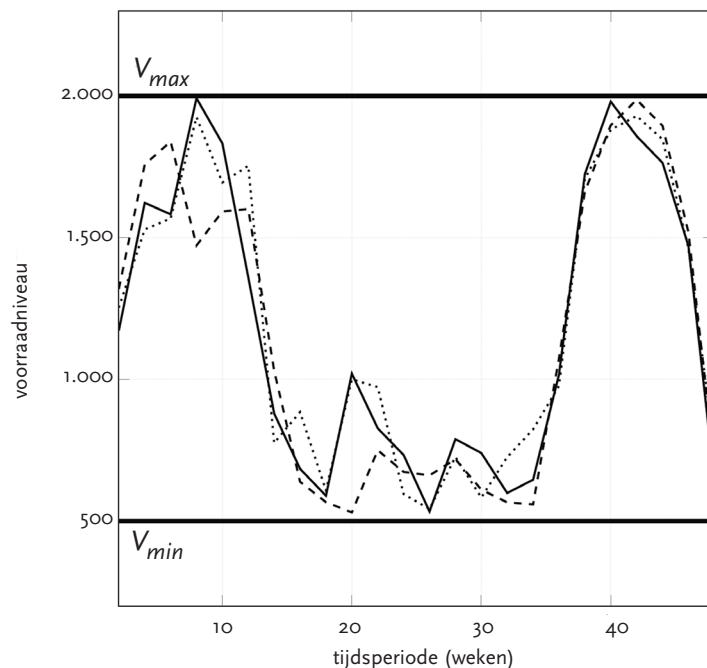
Stel we zijn naïef en nemen aan dat de verkregen data voor de beslisregels exact zijn. We bepalen vervolgens met bestaande ARO-technieken een dynamisch productieplan. We hebben dan een werkend productieplan, maar dat werkt alleen als alle verkregen data exact zijn. In het onderstaande voorbeeld laten we zien dat de oplossingen niet bruikbaar zijn voor het management als er inexacte data gebruikt worden in de beslisregels. Ter illustratie hebben we in ons voorbeeld verlate rapporten over de vraagcijfers als oorzaak van inexacte data. Volledige informatie is dan pas laat beschikbaar; op het moment dat de beslissing genomen moet worden is er alleen een schatting van de werkelijke vraag. Als in periode t slechts 9 van de 10 winkels de vraagcijfers hebben gerapporteerd voor een bepaalde periode, dan is er nog steeds een bepaalde onzekerheid en dus zijn de verkregen data inexact. Resultaten in De Ruiter et al. (2014) laten zien dat de beslisregels niet aan alle restricties voldoen als naïef wordt aangenomen dat de verkregen data exact zijn. Zo kan de opslagcapaciteit overschreden worden of kunnen

we onder het gewenste minimum voorraadniveau uitkomen. In figuur 2 zijn de gevolgen op het voorraadniveau zichtbaar voor drie potentiële scenario's uit ons onzekerheidsgebied. In totaal hebben we 100 verschillende vraagpatronen gesimuleerd en trad het voorraadniveau tot in 80% van de gevallen buiten de gewenste grenzen.

Beslisregels gebaseerd op inexacte data

Het negeren van inexacte data kan dus ernstige gevolgen hebben. Een mogelijkheid is om alle data die inexact zijn niet mee te nemen in ons model. In dat geval gooien we echter waardevolle informatie weg. Het is zelfs zo, dat zonder de informatie in deze data we niet eens een plan kunnen opstellen dat aan alle restricties voldoet. We willen daarom in ons voorraadmodel wel inexacte data gebruiken, maar tegelijkertijd ook rekening houden met het feit dat de verkregen data inexact zijn.

In De Ruiter et al. (2014) is een methode ontwikkeld die als uitgangspunt heeft dat de beslisregels gebaseerd zijn op schatters in plaats van de werkelijke waarden. Voor de afwijking tussen de schatter en de werkelijke waarde van de vraag stellen we wederom een onzekerheidsgebied op. Zo kunnen we bijvoorbeeld het effect bestuderen als er in een bepaalde periode nog 10% onzekerheid is over onze data. Dit betekent dat de werkelijke vraag tot 10% meer of minder kan zijn dan weergegeven wordt door de beschikbare data. We krijgen nu een geheel nieuw onzekerheidsgebied dat zowel parameters, de schatters en de afwijking van de schatters bevat.



Figuur 3. Voorraadniveaus met beslissingsregels gebaseerd op inexacte data: geen overschrijdingen van grenzen.

In principe zijn we weer terug bij een bekend ARO-model, zij het met een nieuw type onzekerheidsgebied. Het fundamentele verschil met het ARO-model is dat beslissingsregels op schatters gebaseerd zijn in plaats van de werkelijke parameterwaarden, die in de praktijk niet geobserveerd kunnen worden. Er kunnen nu verschillende technieken gebruikt worden om dit probleem weer om te schrijven tot een efficiënt oplosbaar probleem. In De Ruiter et al. (2014) wordt gebruik gemaakt van de zogeheten Fenchel dualiteit theorie, die recentelijk toegepast is op robuuste optimalisatie door Ben-Tal et al (2014). Het resulterende probleem wordt weer een lineair programmeringsprobleem dat we efficiënt op kunnen lossen met simplex- of inwendige punt methoden.

Het resulterende productieplan blijft nu binnen de grenzen voor het voorraadniveau blijft, zoals te zien is in figuur 3.

Conclusie

Bij het maken van beslissingen moeten we er rekening mee houden dat verkregen data inexact kunnen zijn. Doen we dit niet, dan kan er aan verschillende restricties in het model niet worden voldaan. Zo kan de opslagcapaciteit overschreden worden of kan het voorraadniveau beneden gewenste minimumwaarden zakken. Voor multi-stage problemen hebben we beslissingsregels gecreëerd die rekening houden met de inexacte aard van data. In De Ruiter (2014) wordt een model beschreven dat dit probleem, met beslissingsregels gebaseerd op inexacte data, efficiënt op kan lossen. Met deze methode kunnen we de inexacte, desalniettemin

waardevolle, data gebruiken om het productieplan op te stellen. Deze nieuwe methode is niet alleen toepasbaar op voorraadproblemen, maar is ook te gebruiken voor andere multi-stage problemen waar verkregen data inexact zijn. Voorbeelden hiervan zijn medische toepassingen en voor de financiële wereld in portfolio management.

LITERATUUR

- Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2014). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2), 351–376.
- Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Oxfordshire: Princeton University Press.
- DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: an empirical analysis. *Management Science*, 54(4), 627–641.
- Haug, A., Zachariassen, F., & Liempd, D. van (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management*, 4(2), 168–193.
- Ruiter, F. J. C. T. de, Ben-Tal, A., Brekelmans, R. C. M., & Hertog, D. den (2014). *Adjustable robust optimizations with decision rules based on inexact revealed data*. (CentER Discussion Paper, 2014-003). Tilburg: Tilburg University, Operations Research.
- Soyster, A. L. (1973). Technical note – convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5), 1154–1157.

FRANS DE RUITER rondde in 2013 zijn master Operations Research aan Tilburg University af en in 2014 zijn master Applicable Mathematics aan de London School of Economics. Hij ontving in 2013 de VVS+OR master thesis prijs voor 'Adjustable robust optimization based on inexact information'. Momenteel is hij PhD student aan Tilburg University waar hij onderzoek doet naar 'Anticipative robust decision making with optimal learning'.
E-mail: <F.J.C.T.deRuiter@uvt.nl>