Tristan van Leeuwen
*Mathematisch Instituut*
*Universiteit Utrecht*
*t.vanleeuwen@uu.nl*

**Column** Tenure-tracker

# Inverse problemen

**In deze rubriek stellen houders van een tenure track-positie zich voor. Om het wiskundig onderwijs en onderzoek op niveau te houden zijn in 2013 tenure track-posities gerealiseerd door de wiskundeclusters. Deze aanstellingen bieden jonge wiskundige onderzoekers kans op aanzienlijke groei in hun wetenschappelijke carrière. Tristan van Leeuwen heeft een tenure track-positie aan de Universiteit van Utrecht.**

Hoe ziet de aarde er van binnen uit? Wat zit er verstopt in een mummie? Hoe diep zit het oliereservoir? We kunnen deze vragen niet beantwoorden met behulp van directe waarnemingen maar hebben vaak wel *indirecte* waarnemingen tot onze beschikking. Voorbeelden zijn aardbevingsmetingen of een CT-scan van een mummie. Als we de invloed van de parameters op de metingen goed kunnen modelleren, kunnen we deze vragen formuleren als een een *invers probleem:*

> Bepaal de parameters, $\mathbf{m}$, waarvoor de voorspelde metingen, $F(\mathbf{m})$, zo goed mogelijk overeenkomen met de daadwerkelijke metingen $\mathbf{d}$.

De voorwaartse operator $F$ beschrijft hoe de parameters de metingen beïnvloeden. Soms kun je een expliciete inverse operator vinden, maar meestal wordt het inverse probleem gesteld als een optimalisatieprobleem:

$$\min_{\mathbf{m}} \rho(F(\mathbf{m}) - \mathbf{d}). \tag{1}$$

De functie $\rho(\cdot)$ meet de fout tussen de voorspelde en gemeten data. Zo nemen we $\rho(\cdot) = \|\cdot\|_2^2$ voor een kleinstekwadratenschatting. In mijn onderzoek richt ik me op verschillende kanten van zulke inverse problemen, met als belangrijkste toepassingen seismische beeldvorming en, sinds kort, computertomografie.

Tijdens mijn promotie ben ik me gaan verdiepen in seismische beeldvorming. De niet-lineariteit van de voorwaartse operator kan er in deze toepassing voor zorgen dat het oplossen van het optimalisatieprobleem (1) niet leidt tot een bruikbare schatting van de parameters. Kort gezegd kan het zijn dat de verkregen oplossing slechts lokaal optimaal is, en niet globaal. Hoewel lokale optimaliteit genoeg is om een optimalisatieprobleem op te lossen, zijn lokaal optimale oplossingen vaak onbruikbaar als oplossing van een inverse probleem. Zulke lokale optima zijn dan een bijverschijnsel van de formulering en door de fout tussen de gemeten en gemodelleerde data op een andere manier te meten, kunnen ze deels worden voorkomen [6].

Als postdoc op de University of British Columbia ben ik me meer gaan richten op oplossingsmethoden voor het optimalistatieprobleem. Ik heb daar onder andere gewerkt aan het automatisch schatten van extra parameters in het inverse probleem. Denk hierbij aan calibratieparameters in $F$ of een covariantiematrix in $\rho$ die bepaalde componenten in de fout meer (of juist minder) gewicht geeft. Het optimalisatieprobleem kan dan worden geschreven als

$$\min_{\mathbf{m},\mathbf{a}} f(\mathbf{m}, \mathbf{a}), \tag{2}$$

waar de vector $\mathbf{a}$ de extra parameters bevat. Hoewel dit probleem als een niet-linear optimalisatieprobleem in $(\mathbf{m}, \mathbf{a})$ kan worden gezien is het vaak efficiënter het probleem te splitsen en een optimalewaardefunctie $\overline{f}(\mathbf{m}) = \min_{\mathbf{a}} f(\mathbf{m}, \mathbf{a})$ te introduceren. Zeker wanneer de optimalisatie in $\mathbf{a}$ makkelijk is, is deze aanpak zeer geschikt. Onder milde voorwaarden is de optimalewaardefunctie differientieerbaar en zijn er zelfs expliciete uitdrukkingen voor de afgeleiden. Zo kunnen we het overgebleven optimalisatieprobleem $\min_{\mathbf{m}} \overline{f}(\mathbf{m})$ oplossen met Newtons methode [2].

Een andere veelvoorkomende structuur die we kunnen uitbuiten is wanneer het optimalisatieprobleem kan worden gesplitst:

$$f(\mathbf{m}) = \sum_{i=1}^{M} f_i(\mathbf{m}). \tag{3}$$

Denk hierbij aan metingen die afkomstig zijn van $M$ onafhankelijke experimenten (verschillende aardbevingen, bijvoorbeeld). Het kan veel rekentijd vergen om een enkele $f_i$ te evalueren, bijvoorbeeld omdat de voorwaartse operator een partiële differentiaalvergelijking (PDV) oplost. Daarnaast is $M$ vaak groot. Om rekentijd te besparen heb ik gewerkt aan dimensiereductiemethoden die de functie $f$ benaderen met een veel kleinere effectieve $M$. Een gerelateerde aanpak om de rekentijd te verminderen is het benaderen van de voorwaartse operator $F$ door bijvoorbeeld de PDV minder nauwkeurig op te lossen. In beide gevallen is het nog steeds de vraag *hoe* nauwkeurig deze benaderingen moeten zijn om te garanderen dat de geschatte parameters, die worden verkregen door de benaderde $f$ te minimaliseren, bruikbaar zijn. Numerieke experimenten wijzen uit dat zulke methoden zeer goede parameterschattingen opleveren in een fractie van de tijd die nodig zou zijn om het gehele probleem met grote nauwkeurigheid door te rekenen [1, 3, 5].

Wanneer de voorwaartse operator een PDV oplost, zoals in seismische beeldvorming, is $F$ impliciet gedefinieerd: $F(\mathbf{m}) = P\mathbf{u}$ waarbij $\mathbf{u}$ de oplossing is van een PDV $A(\mathbf{m})\mathbf{u} = \mathbf{q}$ en $P$ bepaalt welk deel van de oplossing we hebben geobserveerd. Het oplossen van het optimalisatieprobleem (1) leidt in dit geval tot de welbekende geadjungeerde aanpak. We kunnen het probleem (1) ook schrijven als

$$\min_{\mathbf{m},\mathbf{u}} \rho(P\mathbf{u} - \mathbf{d}) \quad \text{zodat} \quad A(\mathbf{m})\mathbf{u} = \mathbf{q}. \tag{4}$$

Een mogelijk voordeel van deze formulering is dat het de zoekruimte vergroot en het daardoor gemakkelijker is een geschikt (globaal) minimum te vinden. Hoewel er elegante methoden bestaan om zulke optimalisatieproblemen met restricties op te lossen via de Lagrangefunctie, zijn deze vaak niet praktisch voor grootschalige problemen omdat de toestandsvariabele $\mathbf{u}$ te groot is. Een interessant alternatief is de zogenaamde penalisatiemethode. We herformuleren het probleem (4) in dit geval als

$$\min_{\mathbf{m},\mathbf{u}} \rho(P\mathbf{u} - \mathbf{d}) + \lambda \|A(\mathbf{m})\mathbf{u} - \mathbf{q}\|_2^2. \tag{5}$$

Nu kunnen we de eerdergenoemde truc toepassen om $\mathbf{u}$ uit het probleem te elimineren. Op deze manier krijgen we een methode die qua rekentijd en geheugengebruik gelijkwaardig is aan de standaard geadjungeerde aanpak, en toch een grotere zoekruimte benut. Een interessante bijkomstigheid van deze aanpak is dat er *geen* expliciete geadjungeerde berekeningen nodig zijn om de benodigde gradiënt uit te rekenen [4]. In de limiet $\lambda \uparrow \infty$, convergeert een oplossing van (5) naar een oplossing van het oorspronkelijke probleem. In toepassingen waar geen grote nauwkeurigheid is vereist, is het vaak voldoende om (5) voor één waarde van $\lambda$ op te lossen.

Na drie jaar in Canada was het tijd om weer eens verder te gaan kijken. Zo kwam ik terecht bij Joost Batenburg op het CWI, met wie ik al sinds mijn verdediging contact had vanwege onze gezamenlijke interesse in beeldvormingstechnieken. Er zit veel overlap in de problematiek bij seismische beeldvorming en computertomografie. Veel van de bovengenoemde technieken zijn dan ook direct toepasbaar in de computertomografie. Daarnaast leidt de specifieke structuur van de problemen in computertomografie tot interessante nieuwe onderzoeksvragen.

Per 1 september 2014 ben ik aan een nieuwe uitdaging begonnen: een tenure track-positie in Utrecht. Ik ben daar van plan mijn onderzoek op het snijvlak van scientific computing en inverse problemen voort te zetten. Daarbij vind ik het leuk de wiskunde aan te wenden om te kijken hoe de aarde er van binnen uitziet, wat er verstopt zit in een mummie, of erachter te komen hoe diep het oliereservoir zit.

**Biografie**

Tristan van Leeuwen werd geboren in 1981 te Hilversum. Hij studeerde computational science in Utrecht en schreef zijn afstudeerscriptie, onder begeleiding van Rob Bisseling, over het partitioneren van hypergrafen. Van 2006–2010 deed hij zijn promotieonderzoek aan de TU Delft, onder begeleiding van Wim Mulder, en ontwikkelde een nieuwe methode om het seismische inverse probleem aan te pakken. Meteen daarna verhuisde hij met zijn vriendin naar Vancouver, Canada. Als postdoc ging hij aan de slag in de groep van Felix Herrmann en werkte daar onder andere aan dimensiereductietechnieken en optimalisatietechnieken en hun toepassing in de seismische beeldvorming. In 2013 verhuisden ze weer terug naar Nederland, waar Tristan als postdoc bij het CWI ging werken. In de groep van Joost Batenburg verdiepte hij zich in computertomografie. Sinds 1 september 2014 is hij begonnen als tenure-tracker bij het Mathematisch Instituut van de Universiteit Utrecht.

## Referenties


1. A.Y. Aravkin, M.P. Friedlander, F.J. Herrmann en T. Leeuwen, Robust inversion, dimensionality reduction, and randomized sampling, *Mathematical Programming* 134 (2012), 101–125.
2. A.Y. Aravkin en T. van Leeuwen, Estimating nuisance parameters in inverse problems, *Inverse Problems* 28 (2012), 115016.
3. T. van Leeuwen en F.J. Herrmann, Fast waveform inversion without source-encoding, *Geophysical Prospecting* 61 (2013), 10–19.
4. T. van Leeuwen en F.J. Herrmann, Mitigating local minima in full-waveform inversion by expanding the search space, *Geophysical Journal International* 195 (2013), 661–667.
5. T. van Leeuwen en F.J. Herrmann, 3D Frequency-domain seismic inversion with controlled sloppiness, te verschijnen in *SIAM Journal on Scientific Computing* (2014).
6. T. van Leeuwen en W.A. Mulder, A correlation-based misfit criterion for wave-equation traveltime tomography, *Geophysical Journal International* 182 (2010), 1383–1394.