Bayesian inference for Gaussian models: Inverse problems and evolution equations

Yan, D.

Author: Yan, D.
Title: Bayesian inference for Gaussian models: Inverse problems and evolution equations
Issue Date: 2020-03-03

Samenvatting

In dit proefschrift wordt de theorie achter het asymptotische gedrag van Bayesiaanse inferentie voor niet-parametrische Gaussische lineaire modellen ontwikkeld. De modellen hebben de vorm van

$$\text{Observaties} = \underset{\text{zoals een gladgestreken parameter}}{\text{Getransformeerde signaal}} + \text{Gaussische ruis},$$

en het doel is om het originelt signaal te reconstrueren met behulp van de Bayesiaanse methoden.

In Deel I beschrijven we de benodigde concepten voor het ontwikkelen van de theorie in dit proefschrift. In het bijzonder richten we ons op de theorie voor het getransformeerde-signaal-in-ruis model. Het doel is om het (originele) signaal terug te schatten met zo min mogelijke aannames op de parameter ruimte, en met een hoge convergentiesnelheid. Aangenomen wordt dat de parameter (signaal) aan bepaalde gladheidsvoorwaarden voldoet, welke gekwantificeerd worden in het eerste gedeelte van dit proefschrift. De ruisstructuur van het statistisch model is ook van belang, omdat het een grote invloed heeft op de moeilijkheid van het schattingsprobleem. Voor het schatten gebruiken we Bayesiaanse methoden, welke formeel gedefinieerd worden in Hoofdstuk 4.

In Hoofdstuk 2 worden gladheidsklassen geïntroduceerd om de mate van regulariteit van de parameterruimte te kwantificeren. Een eerste gladheidsklasse die we bespreken zijn zogeheten gladheidsschalen. Een gladheidsschaal $\{H_s\}_{s\in\mathbb{R}}$ is een collectie van Hilbertruimtes H_s die op basis van de normen is genest, en welke voldoet aan een zogeheten norm dualiteit. Een bijzonder subklasse van gladheidsschalen zijn Hilbertschalen. Een Hilberschaal is een gladheidsschaal waarbij de index s op natuurlijke wijze een genererende operator definieert. Met de resulterende genererende operator kan er een link gelegd worden tussen de Hilbertschaal en de covariantie operator van Gaussische a priori verdelingen en Gaussische ruis. Een ander type gladheidsklasse dat aan bod komt in dit hoofdstuk wordt gebruikt om de anisotropische gladheid in hogere dimensies te beschrijven. Deze laatste gladheidsklasse wordt gebruikt om stochastische evolutie vergelijkingen te bestuderen in Hoofdstuk III. We definiëren ook benaderingsgetallen die de benaderingseigenschappen van gladheidsklassen beschrijven, en we relateren deze aan metrische entropiën.

In Hoofdstuk 3 geven we een formele beschrijving van Gaussische maten in oneindig dimensionale ruimtes. Eerst geven we een overzicht van kansmaten op

Banachruimtes in het algemeen. Nadat we Gaussische maten formeel hebben gedefinieerd, laten we zien dat ze gunstige eigenschappen hebben. Van bijzonder belang is de covariantie structuur van Gaussische maten, wat we benadrukken met een aantal voorbeelden. Daarnaast geven we ook een kort overzicht van cylindrische maten en radonificatie in dit hoofdstuk.

In Hoofdstuk 4 geven we een formele beschrijving van Bayesiaanse inferentie in oneindig dimensionale ruimtes. Hiervoor introduceren we de regel van Bayes voor algemene functieruimtes, en we bespreken de basisconcepten van de asymptotiek zoals consistentie en convergentiesnelheid. De regel van Bayes leidt tot een natuurlijke schattingsprocedure, welke we toepassen op Gaussische lineaire models. We zullen in het bijzonder geïnteresseerd zijn in het geval met continue observaties, en het geval met discrete observaties. Het hoofdstuk wordt afgesloten met een algemene stelling over a posteriori convergentiesnelheden met een algemene bewijsmethode welke de basis vormt voor soortgelijke resultaten in de rest van dit proefschrift.

In Deel II richten we op lineaire inverse problemen met Gaussische ruis, waarvoor de theorie van Gaussische lineaire modellen toepasbaar is. In dit geval observeren we een gladgestreken signaal dat verontreinigd is door ruis. Door het gladstrijken van het signaal wordt het schatten bemoeilijkt. De naïeve benadering resulteert in slechte schattingen van het origineel signaal, doordat de ruis opgeblazen wordt door de inverse operator. Een oplossing is om de inverse operator te regulariseren, wat op natuurlijke wijze gedaan kan worden met de Bayesiaanse methode. In dit gedeelte van de proefschrift bestuderen we inverse problemen voor twee observatieregimes: Continu en discreet geobserveerde inverse problemen die ook bekend staan als het witte ruis model en regressie.

In Hoofdstuk 5 worden lineaire inverse problemen heuristisch en wiskundig rigoureus beschreven. Om het geheel tastbaar te maken geven we ook enkele voorbeelden. In dit hoofdstuk wordt ook de relatief algemene projectie van Galerkin belicht, omdat we hiermee vrij gemakkelijk convergentiesnelheden voor inverse problemen kunnen afleiden zoals beschreven is in de rest van Deel II.

In Hoofdstuk 6 beschrijven we het asymptotisch gedrag van niet-parametrische Bayesiaanse methoden toegepast op lineaire inverse problemen met continue observaties. Een algemene stelling voor continu geobserveerde inverse problemen wordt bewezen door een aanpassing in het bewijs van de algemene stelling over convergentiesnelheden dat afgeleid is in Hoofdstuk 4. In het algemeen is adaptiviteit niet gegarandeerd, want daar zijn extra condities voor nodig welke afgeleid en beschreven zijn in de tweede stelling in dit hoofdstuk. De eerste stelling wordt gebruikt om het asymptotisch gedrag van twee Bayesiaanse procedures te bestuderen: De procedure op basis van a priori verdelingen die beschreven zijn als gerandomiseerde reeksen, en de procedure op basis van Gaussische a priori verdelingen. De a priori verdelingen die beschreven worden als gerandomiseerde reeksen resulteren in a posteriori verdelingen die, op een vermenigvuldiging met een logaritmische factor na, convergeren met de minimax convergentiesnelheid. De Gaussische a priori verdelingen leiden echter tot a posteriori verdelingen die met een suboptimale convergentiesnelheid krimpen. Om het laatste geval te verbeteren introduceren wij mengvormen van Gaussische a priori verdelingen, door een hogere orde a priori verdeling te kiezen op de Gaussische a priori verdelingen, wat resulteert in een procedure die adaptief en optimaal is in de minimax zin. Dit laatste wordt bewezen met behulp van de tweede stelling van dit hoofdstuk.

In Hoofdstuk 7 beschrijven we onze eerste aanpak van discreet geobserveerde inverse problemen. Dit doen we door gebruik te maken van de Gaussische conjugatie eigenschap in lineaire problemen. Hier wordt “discreet” gebruikt om aan te geven dat de observaties alleen op discrete punten zijn gemeten, terwijl in het continu geval het gehele pad geobserveerd is. Een reeksformulering van de discreet geobserveerde inverse problemen is afgeleid op basis van een singulierewaardenontbinding van de gladstrijkende operator. De hoofdresultaten beschrijven de convergentiesnelheden en de a posteriori plausibele verzamelingen voor milde en extreem singuliere inverse problemen. De resultaten worden geïllustreerd met simulaties.

In Hoofdstuk 7 beschrijven we onze tweede aanpak van discreet geobserveerde inverse problemen. In dit hoofdstuk gebruiken we de resultaten van Hoofdstuk 6 om het model te bestuderen. Hiervoor gebruiken we een signaalreconstructietechniek dat continue paden construeert door de discrete punten te interpoleren, waarvoor de interpolatiefout ook geschat kan worden. We veralgemeniseren dan de stelling van Hoofdstuk 6, welke alleen geldt voor inverse problemen met continue observaties, zodat het ook toepasbaar is op discrete geobserveerde inverse problemen. Dit resulteert in convergentiesnelheden voor a priori verdelingen die beschreven worden als gerandomiseerde reeksen, Gaussische a priori verdelingen, en mengvormen van Gaussische a priori verdelingen.

In Deel III richten we ons op niet-parametrisch Bayesiaanse inferentie voor stochastische evolutie vergelijkingen. Er wordt aangenomen dat de toestandsruimte van de stochastische evolutie vergelijking oneindig dimensionaal is, en dat de bijbehorende dynamische systeem beïnvloed wordt door een ruisproces. In dit gedeelte wordt er van uitgegaan dat de dynamica gedreven wordt door een deterministische component die beschreven wordt door een lineaire partiële differentiaalvergelijking, en een stochastische component die beschreven wordt door een additieve Gaussische ruis over zowel de toestandsruimte als de tijdsdimensie. Dit model kan gezien worden als een bevordering van het witte ruis model naar een oneindig dimensionale toestandsruimte. Het doel in dit gedeelte van de proefschrift is om theorie te ontwikkelen voor Bayesiaanse inferentie voor de driftterm en de beginwaarden van de evolutie vergelijkingen.

In Hoofdstuk 9 geven we een algemene overzicht van stochastische evolutie vergelijkingen. Eerst definiëren we $\mathcal{Q}$-Wienerprocessen, een veralgemenisering van Brownse beweging naar oneindig dimensionale ruimtes. Met deze $\mathcal{Q}$-Wienerprocessen kunnen we een klasse van stochastische integralen over Hilbertruimtes beschrijven, welke daarna ook veralgemeniseerd wordt zodat het toepasbaar is op een grotere klasse van integranden. Deze integralen gecombineerd met de theorie van deterministische evolutie vergelijkingen worden dan gebruikt om stochastische evolutie vergelijkingen formeel te definiëren. We laten ook zien dat de oplossingen van deze vergelijkingen gerepresenteerd kunnen worden als stochastische processen.

In Hoofdstuk 10 beschrijven we het asymptotisch gedrag van de Bayesiaanse schattingsprocedure voor de beginwaarden en de driftterm van de stochastische evolutie vergelijkingen op basis van Gaussische a priori verdelingen. Om convergentiesnelheden af te leiden passen we de algemene aanpak die beschreven is in Hoofdstuk 4 aan. Om de beginwaarden te schatten construeren we een type spatiële Gaussische a priori verdelingen, waarvoor we ook de snelheid kunnen afleiden waarmee de a posteriori verdelingen zich concentreert om de ware beginwaarden. Dit schattingsprobleem kan gezien worden als een veralgemenisering van het extreme singuliere inverse probleem dat beschreven wordt in Hoofdstuk 7. Voor de Bayesiaanse schattingsprocedure van de driftterm construeren we een ander type Gaussische a priori verdelingen. Deze a priori verdelingen zijn ontwikkeld om de anisotropische gladheid van signalen in hoger dimensies te identificeren. Een cruciaal punt is dat we de observaties dusdanig kunnen transformeren zodat het model geconverteerd wordt in een multi-index Gaussische reeks model. Met dit inzicht kunnen we de convergentiesnelheden afleiden.