Quellentrennung in hydroakustischen Zeitreihen als Inverses Problem

Andreas Galka

Wehrtechnische Dienststelle für Schiffe und Marinewaffen, Maritime Technologie und Forschung (WTD 71), 24340 Eckernförde, Deutschland; E-Mail: andreasgalka@bundeswehr.org

Einleitung

Hydroakustische Zeitreihen enthalten oft Gemische verschiedener Signale natürlichen oder technischen Ursprungs. Im Allgemeinen sind sowohl die Mischungsfaktoren als auch die Eigenschaften der Quellsignale zunächst unbekannt. Bei Verwendung eines Einzelhydrophons entfällt die Möglichkeit, mittels Beamforming die Quellsignale räumlich zu trennen.

In dieser Situation können Algorithmen zur Lösung Inverser Probleme, wie die Zustandsraum-Modellierung mittels *Maximum-Likelihood*, helfen. Die numerische Maximierung, bei Vorliegen von Nebenbedingungen, kann mittels des *Expectation-Maximisation*-Algorithmus, sowie weiterer Verfahren, implementiert werden.

Aufbauend auf [1], wird in diesem Paper ein Algorithmus zur Rekonstruktion von Quellsignalen aus einem linearen Gemisch weiter entwickelt, und es werden Ergebnisse für reale Zeitreihen von Einzelhydrophonen vorgestellt.

Die Entwicklung dieser Algorithmen wird motiviert durch die Aufgabe, im Wasserschall Quellsignale zu detektieren, zu trennen und zu klassifizieren. Anwendungsgebiete sind die Untersuchung von Schiffssignaturen, akustische Unterwasserkommunikation, die Detektion transienter Ereignisse, aber auch die Untersuchung biologischer Quellen, wie etwa Meeressäuger.

Inverse Probleme

Wir gehen davon aus, daß Daten in Gestalt n-dimensionaler Vektoren $\mathbf{y}(t)$ vorliegen, die von einem dynamischen System erzeugt wurden, dessen interne Zustände durch einen zeitabhängigen m-dimensionalen Zustandsvektor $\mathbf{x}(t)$ beschrieben werden; die Komponenten $x_i(t)$ von $\mathbf{x}(t)$ stellen die gesuchten Quellsignale dar. Der Zusammenhang zwischen $\mathbf{x}(t)$ und $\mathbf{y}(t)$ wird dann beschrieben durch eine Meßgleichung

$$\mathbf{y}(t) = g\big(\mathbf{x}(t), \boldsymbol{\epsilon}(t)\big)$$

wobei $g(.)$ eine (i.Allg. nichtlineare) Meßfunktion darstellt und $\boldsymbol{\epsilon}(t)$ einen Störterm repräsentiert, der aus Meßrauschen bestehen kann. Das Vorwärts-Problem besteht dann darin, für gegebenes $\mathbf{x}(t)$ den Datenvektor $\mathbf{y}(t)$ zu berechnen – was zumeist keine Probleme bereitet – , während das Inverse Problem darin besteht, für gegebenes $\mathbf{y}(t)$ den Zustandsvektor $\mathbf{x}(t)$ zu bestimmen.

Wenn die Meßgleichung als linear mit additivem Meßrauschen (dessen Kovarianzmatrix mit R bezeichnet werden soll) angenommen wird, nimmt sie folgende Form an:

$$\mathbf{y}(t) = \mathsf{C}\mathbf{x}(t) + \boldsymbol{\epsilon}(t) \qquad (1)$$

wobei C eine $(n \times m)$-dimensionale Beobachtungsmatrix ist. Das Inverse Problem besteht dann also im Wesentlichen darin, C zu invertieren. Dies ist aber nicht ohne weiteres möglich, da C zumeist nicht quadratisch ist und oft sogar unbekannt.

Zustandsraum-Modellierung

Um Inverse Probleme eindeutig lösen zu können, werden in der Multivariaten Statistik oft zusätzliche Annahmen gemacht. Klassische Beispiele sind:

- $n > m$; $x_i(t)$ gauß-verteilt, jeweils mit Varianz 1.0 und paarweise *unkorreliert*; R diagonal; $\mathsf{C}^\intercal\mathsf{R}^{-1}\mathsf{C}$ diagonal. Diese Annahmen führen zur *Faktorenanalyse* (FA).
- $n \geq m$; $x_i(t)$ nicht gauß-verteilt[1] jeweils mit Varianz 1.0 und paarweise *unabhängig*; $\mathsf{R} = \mathsf{0}$. Diese Annahmen führen zur *Independent Component Analysis* (ICA).

Eine Alternative zu FA und ICA ist die (lineare) Zustandsraum-Modellierung (*state space modelling*) [2], die auf folgenden Annahmen beruht:

- $\mathbf{x}(t)$ folgt einer linearen Dynamik:

$$\mathbf{x}(t) = \mathsf{A}\mathbf{x}(t-1) + \boldsymbol{\eta}(t) \qquad (2)$$

  worin A eine $(m \times m)$-dimensionale Zustandsübergangsmatrix und $\boldsymbol{\eta}(t)$ ein m-dimensionaler gauß-verteilter Rauschterm ist, dessen Kovarianzmatrix mit Q bezeichnet werden soll.

Unter der Voraussetzung, daß A und C eine bestimmte Bedingung erfüllen (*observability*), kann die Zustandsraum-Modellierung auch im unterbestimmten Fall $n < m$ die Sequenz der Zustandvektoren $\mathbf{x}(t)$ rekonstruieren, d.h., das Inverse Problem lösen.

In diesem Paper zeigen wir Beispiele für den Extremfall $n = 1$, d.h., aus den Daten eines einzelnen Hydrophons sollen eine Anzahl von Quellsignalen rekonstruiert werden. Dieses Inverse Problem kann nur durch Zustandsraum-Modellierung gelöst werden, nicht aber durch FA oder ICA.

Die Rekonstruktion von $\mathbf{x}(t)$ erfolgt für gegebene Daten $\mathbf{y}(t)$ und bekannte Parametermatrizen A, C, Q und R durch *Kalman-Filterung*; dieser Algorithmus entspricht in etwa einer regularisierten Pseudo-Invertierung von C. Da die Parameter-Matrizen aber zunächst

[1] Maximal eine Komponente von $\mathbf{x}(t)$ darf gauß-verteilt sein.

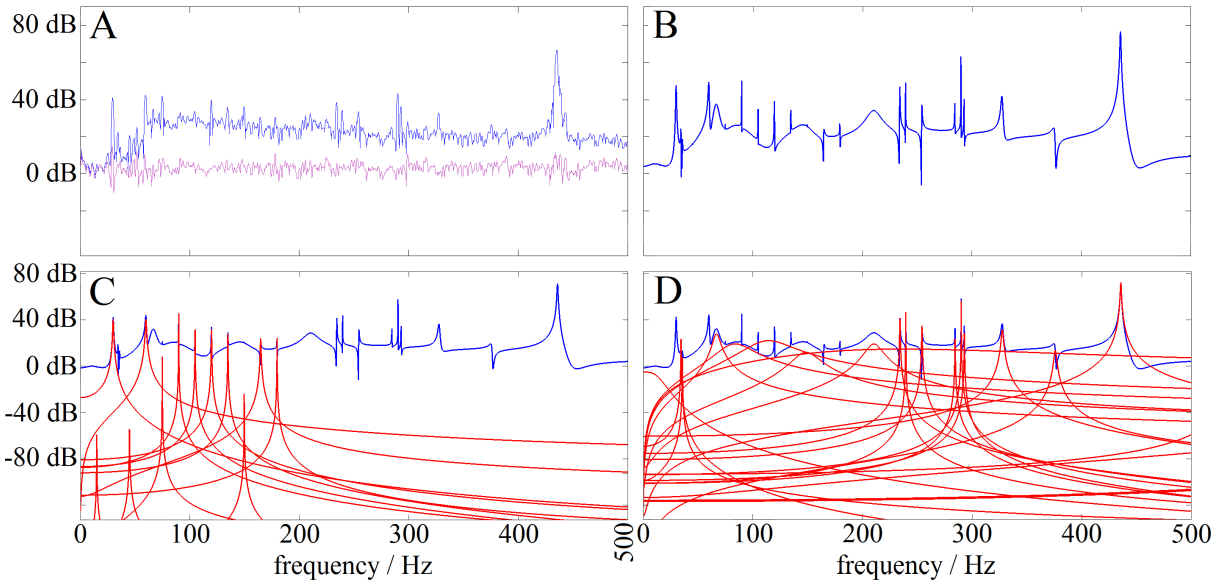

Abbildung 1: Analyse von Wasserschall mit Schiffslinien. A: Nicht-parametrisches Leistungsspektrum (DFT) der Daten (blau) und der Vorhersagefehler (lila). B: Parametrisches Leistungsspektrum der Daten. C: Beiträge der Linien der Diesel-Maschine (rot) zum parametrischen Leistungsspektrum (blau). D: Andere Beiträge (rot) zum parametrischen Leistungsspektrum (blau).

nicht bekannt sind, werden sie durch einen numerischen *Maximum-Likelihood*-Ansatz abgeschätzt; dies erfolgt zum Teil durch den von Shumway und Stoffer [3] für Zustandsraum-Modelle eingeführten *Expectation-Maximisation*-Algorithmus, zum Teil durch Standard-Verfahren der numerischen Optimierung, wie den Nelder-Mead-Simplex-Algorithmus und den Broyden-Fletcher-Goldfarb-Shanno-Algorithmus [4].

Dynamische Modelle

Die Dynamik der Quellsignale wird durch ARMA-Prozesse modelliert, wobei ARMA für *autoregressive moving-average* steht:

$$x_i(t) = \sum_{\tau=1}^{p} a_i(\tau) x_i(t-\tau) + \sum_{\tau=1}^{p-1} b_i(\tau)\eta_i(t-\tau) + \eta_i(t)$$

Hier repräsentiert der ganzzahlige Parameter p die Modellordnung. Meistens wird $p = 1$ oder $p = 2$ gewählt:

$$x_i(t) = a_i(1)x_i(t-1) + \eta_i(t)$$
$$x_i(t) = a_i(1)x_i(t-1) + a_i(2)x_i(t-2) + b_i(1)\eta_i(t-1) + \eta_i(t)$$

Diese Modelle werden mit AR(1), bzw. ARMA(2,1) bezeichnet. Die autoregressiven Parameter $a_i(1), a_i(2)$ entsprechen einer Frequenz ϕ_i und einem Dämpfungsparameter ρ_i gemäß

$$\rho_i \exp(j\phi_i) = \frac{1}{2}\left(a_i(1) + \sqrt{[a_i(1)]^2 + 4a_i(2)}\right) \quad (3)$$

wobei $j := \sqrt{-1}$. ARMA(2,1)-Modelle werden in ein Zustandsraum-Modell überführt gemäß

$$\begin{pmatrix} x_i(t) \\ \tilde{x}_i(t) \end{pmatrix} = \underbrace{\begin{pmatrix} a_i(1) & 1 \\ a_i(2) & 0 \end{pmatrix}}_{\mathsf{A}_i} \begin{pmatrix} x_i(t-1) \\ \tilde{x}_i(t-1) \end{pmatrix} + \begin{pmatrix} 1 \\ b_i(1) \end{pmatrix} \eta_i(t)$$

wobei $\tilde{x}_i(t)$ eine zusätzliche Zustandskomponente ist, die sich nicht direkt auf die Daten auswirkt; im allgemeinen Fall sind $p-1$ solche Komponenten notwendig. Die Kovarianzmatrix des Rauschterms ergibt sich als

$$\mathsf{Q}_i = \begin{pmatrix} 1 & b_i(1) \\ b_i(2) & [b_i(1)]^2 \end{pmatrix}$$

wobei die Varianz von $\eta_i(t)$ zu 1.0 normiert wurde. Für eine Anzahl N von Quellsignalen, die mittels ARMA(2,1)-Prozessen modelliert werden sollen, ergeben sich die Matrizen A und Q in Gleichung (2), indem jeweils für $i = 1, \ldots, N$ die A_i und die Q_i zu Block-Diagonal-Matrizen angeordnet werden. Diese Struktur bewirkt die paarweise Unabhängigkeit der einzelnen ARMA-Prozesse.

Die Beobachtungsmatrix C enthält für jede ARMA(2,1)-Komponente einen Amplitudenparameter c_i sowie eine Null für die zusätzliche Zustandskomponente. Schätzwerte für die Parameter $a_i(1)$, $a_i(2)$, $b_i(1)$ und c_i müssen durch Maximierung der Likelihood bestimmt werden.

Parametrische Leistungsspektren

Zu jedem ARMA$(p, p-1)$-Modell gehört ein parametrischer Schätzer des Leistungsspektrums, gegeben durch

$$\tilde{y}_i(f) = \left| \frac{c_i\left(1 + \sum_{k=1}^{p-1} b_i(k)\exp(-2k\pi jf)\right)}{1 - \sum_{k=1}^{p} a_i(k)\exp(-2k\pi jf)} \right|^2 \quad (4)$$

Im Fall $p = 2$ ergibt sich

$$\tilde{y}_i(f) = \left| \frac{c_i\left(1 + b_i(1)\exp(-2\pi jf)\right)}{1 - a_i(1)\exp(-2\pi jf) - a_i(2)\exp(-4\pi jf)} \right|^2 \quad (5)$$

Der Fall $p = 1$ folgt entsprechend. Das Gesamtspektrum von N ARMA(2,1)-Modellen ergibt sich zu

$$\tilde{y}(f) = \left| \sum_{i=1}^{N} \frac{c_i\left(1 + b_i(1)\exp(-2\pi jf)\right)}{1 - a_i(1)\exp(-2\pi jf) - a_i(2)\exp(-4\pi jf)} \right|^2 \quad (6)$$

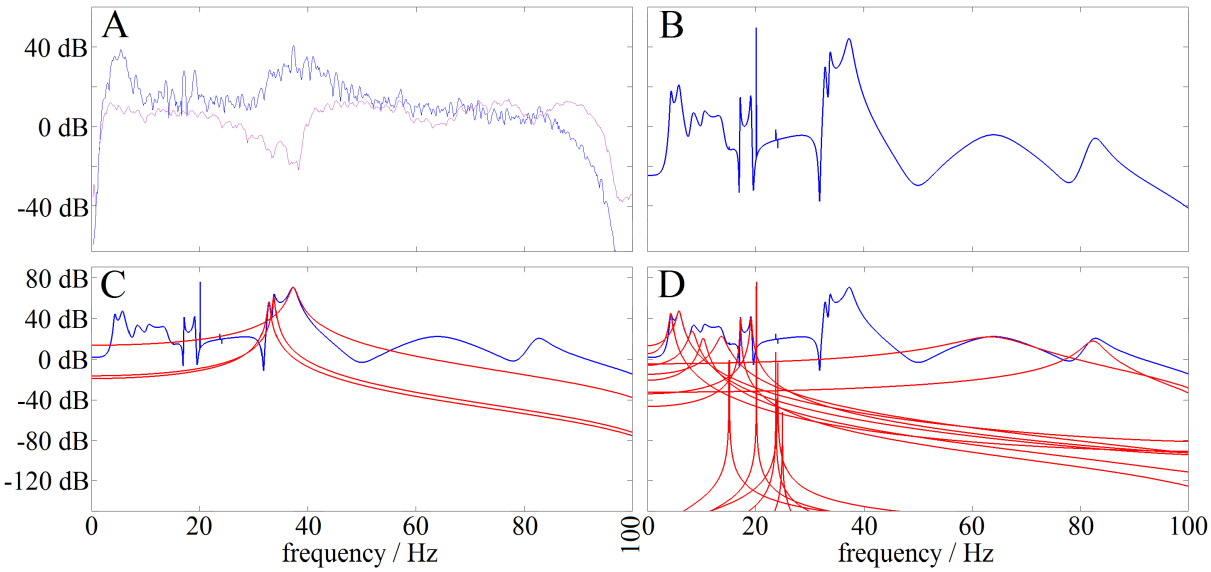

Abbildung 2: Analyse von Wasserschall mit Stimmsignalen eines Blauwals. A: Nicht-parametrisches Leistungsspektrum (DFT) der Daten (blau) und der Vorhersagefehler (lila). B: Parametrisches Leistungsspektrum der Daten. C: Beiträge der Komponenten des Blauwal-Stimmsignals (rot) zum parametrischen Leistungsspektrum (blau). D: Andere Beiträge (rot) zum parametrischen Leistungsspektrum (blau).

Anwendungsbeispiel: Schiffslinien

Als erstes Anwendungsbeispiel wird ein Datensatz ausgewählt, der im August 2019 bei einer Vorbeifahrt des Forschungsschiffs ELISABETH MANN BORGESE an einer abgesetzten AMAR-Boje (JASCO Applied Sciences) in der Ostsee im Flachwasser einige Seemeilen vor der Einfahrt des Hafens Warnemünde aufgezeichnet wurde; die Geschwindigkeit betrug 5.6 kn. Die Abtastrate beträgt, nach Subsampling, 2 kHz. Für die Analyse wird ein Intervall von 20 Sekunden Länge ausgewählt.

Das nicht-parametrische Leistungsspektrum der Daten ist in Abb. 1, Panel A, gezeigt (blau); man erkennt eine größere Anzahl von Schiffslinien.

Die Modellierung erfolgt durch N=34 ARMA(2,1)-Modelle; dabei wird bei der Modellierung angestrebt, daß diese Modelle sämtliche zeitliche Korrelation erfassen (insbesondere die schmalbandigen Schiffslinien), so daß die Vorhersagefehler des Modells nur noch aus weißem Rauschen bestehen. Das nicht-parametrische Leistungsspektrum der Vorhersagefehler ist ebenfalls in Abb. 1, Panel A, gezeigt (lila). Es ist erkennbar, daß das Ziel eines weißen Spektrums nahezu erreicht ist.

Das parametrische Leistungsspektrum der Daten gemäß Gleichung (6) ist in Abb. 1, Panel B, gezeigt; deutlich sind die erfaßten Schiffslinien zu erkennen, aber auch breitbandige Anteile. Eine Reihe von Schiffslinien sind Vielfache einer Grundfrequenz von 15.13 Hz; sie stellen den Frequenzkamm der Diesel-Maschine der ELISABETH MANN BORGESE dar. In Abb. 1, Panel C, sind die Beiträge von 12 ARMA(2,1)-Modellen (rot), die diesen Frequenzkamm repräsentieren, zum parametrischen Leistungsspektrum der Daten (blau) gezeigt, gemäß Gleichung (5); in Panel D sind die Beiträge der übrigen 22 ARMA(2,1)-Modelle gezeigt.

Während der numerischen Optimierung wurden die Frequenzen der Komponenten des Frequenzkamms der Diesel-Maschine miteinander gekoppelt, so daß nur eine Frequenzvariable angepaßt werden mußte.

Als weiteres Ergebnis neben den parametrischen Leistungsspektren liefert die Zustandsraum-Modellierung auch den zeitlichen Verlauf der Komponenten, was bei Vorliegen von Nichtstationarität nützlich sein kann.

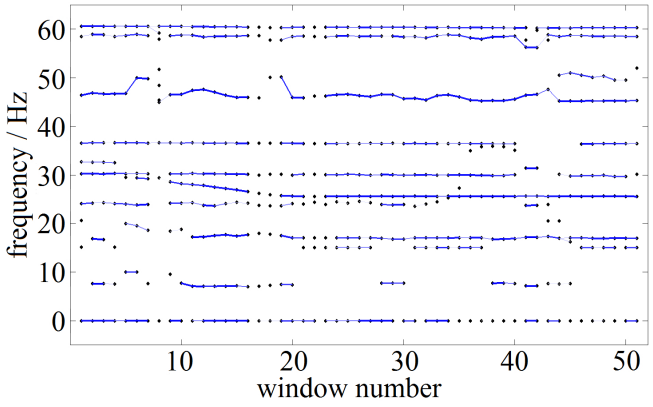

Abbildung 3: Analyse von Wasserschall mit Schiffslinien: Frequenzen der Komponenten von Zustandsraum-Modellen, berechnet für 51 Datenfenster von je 6 Sekunden Länge, mit 50% Überlapp. Jedes Fenster wurde durch 28 ARMA(2,1)-Modelle sowie ein AR(1)-Modell modelliert, wobei jeweils das Zustandsraum-Modell des vorangegangenen Fenster zur Initialisierung der *Maximum-Likelihood*-Optimierung verwendet wurde. Aufgrund des Überlapps können die Korrelationen zwischen den Zustandskomponenten in den jeweils doppelt modellierten 3-Sekunden-Intervallen analysiert werden, so daß die Komponenten in den Einzelmodellen über die Sequenz von Datenfenstern hinweg verfolgt und durch Linien verbunden werden können (blau). Linien sind nur für Werte der Korrelation über einem Schwellwert von 0.75 dargestellt.

Bei längeren Datensätzen kann Nichtstationarität der Frequenzen durch wiederholte Modellierung auf einem sich verschiebenden Fenster modelliert werden. Ein Beispiel einer solchen Analyse, wiederum für eine AMAR-Hydrophon-Messung der ELISABETH MANN BORGESE, ist in Abb. 3 gezeigt. In der Abbildung ist u.a. eine Komponente sichtbar, deren Frequenz innerhalb von 30 Sekunden von 29.3 Hz auf 26.6 Hz absinkt; es handelt sich offenbar um eine Schiffslinie mit (vermutlich technisch bedingt) abfallender Frequenz. Abb. 3 zeigt somit ein *parametrisches Lofargramm*.

Anwendungsbeispiel: Meeressäuger

Als zweites Anwendungsbeispiel wird ein Datensatz ausgewählt, der im Atlantik aufgenommen wurde und Infraschall-Stimmsignale von Blauwalen enthält; es sind auch einige schwache Schiffslinien vorhanden (der Datensatz wurde veröffentlicht in [5]). Die Abtastrate beträgt, nach Subsampling, 220.5 Hz; bei der Messung wurden durch Tiefpaß-Filterung Signalanteile oberhalb von etwa 90 Hz unterdrückt. Für die Analyse wird zunächst ein Intervall von 56 Sekunden Länge ausgewählt.

Das nicht-parametrische Leistungsspektrum der Daten ist in Abb. 2, Panel A, gezeigt (blau); man erkennt ein breitbandiges Spektrum mit einem relativ breiten Peak um 40 Hz, der dem Blauwal-Stimmsignals entspricht, sowie einige schwache schmalbandige Linien um 20 Hz.

Die Modellierung erfolgt durch N=18 ARMA(2,1)-Modelle; das nicht-parametrische Leistungsspektrum der Vorhersagefehler ist ebenfalls in Abb. 2, Panel A, gezeigt (lila). Das Ziel eines weißen Spektrum wird annähernd erreicht, bis auf einen Einbruch um 40 Hz herum; dies ist möglicherweise ein Effekt des Tiefpaß-Filters.

Das parametrische Leistungsspektrum der Daten gemäß Gleichung (5) ist in Abb. 2, Panel B, gezeigt; es ist zu erkennen, daß sowohl breitbandige, als auch schmalbandige Anteile auftreten. In Abb. 2, Panel C, sind die Beiträge von 3 ARMA(2,1)-Modellen (rot), die das Blauwal-Stimmsignal repräsentieren, zum parametrischen Leistungsspektrum der Daten (blau) gezeigt, gemäß Gleichung (4); die zugehörigen Frequenzen sind 33.11 Hz, 34.55 Hz und 37.18 Hz. In Panel D sind die Beiträge der übrigen 15 ARMA(2,1)-Modelle gezeigt; um 20 Hz herum erkennt man 7 schmalbandige Komponenten, die vermutlich Schiffslinien repräsentieren.

Auch für diesen Datensatz liefert die Zustandsraum-Modellierung den zeitlichen Verlauf der Komponenten; in Abb. 4 wird dies anhand der drei Komponenten, die das Blauwal-Stimmsignal repräsentieren, illustriert, dabei werden zum Vergleich auch die ursprünglichen Daten gezeigt (blau). Diesmal wurde das Zustandsraum-Modell auf ein Intervall von 206 Sekunden Länge angewendet, obwohl es nur unter Verwendung der ersten 56 Sekunden erstellt worden war, d.h., der Kalman-Filter wurde auf die restlichen 150 Sekunden *out of sample* angewendet. Es ist deutlich sichtbar, daß in dem gewählten Intervall 5 separate Stimmäußerungen des Blauwals vorliegen; im Verlaufe der numerischen Maximierung der Likelihood wurden diese mittels dreier Komponenten erfaßt, jede mit einer eigenen Peak-Frequenz (siehe Abbildung), so daß auch die zeitliche Veränderung der Stimmäußerungen bezüglich der Frequenz modelliert wird, in dieseem Fall ohne ein sich verschiebendes Fenster.

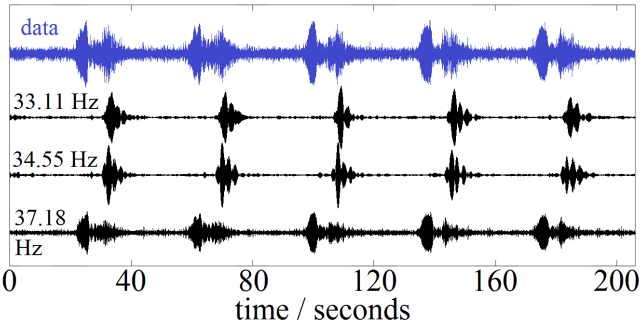

Abbildung 4: Analyse von Wasserschall mit Stimmsignalen eines Blauwals: Daten (blau) sowie 3 von insgesamt 18 Komponenten im Zustandsraum (schwarz), mit jeweiliger Peak-Frequenz gemäß Gleichung (3).

Zusammenfassung

In diesem Beitrag ist gezeigt worden, daß das Inverse Problem der Quellentrennung (*blind signal separation*) durch lineare Zustandsraum-Modellierung gelöst werden kann, auch im unterdeterminierten Fall $n < m$. Zu diesem Zweck werden die Quellsignale als stochastische lineare Oszillationen mit konstanter Frequenz modelliert.

Die Modellierung durch ARMA-Prozesse im Zustandsraum ermöglicht die Berechnung *parametrischer Leistungsspektren* und *parametrischer Lofargramme*.

Auch inhärent nichtstationäre Signale, wie die Blauwal-Stimmäußerungen, wurden korrekt detektiert, obwohl das verwendete Zustandsraum-Modell Nichtstationarität nicht explizit berücksichtigt.

Als Nachteil der vorgestellten Methode ist der hohe Zeitaufwand für die numerische Maximierung der Likelihood für Zustandsraum-Modelle zu erwähnen. In zukünftiger Arbeit sollen effiziente Algorithmen entwickelt werden, die diesen Zeitaufwand reduzieren.

Literatur

[1] Galka, A.: Parametrische Kurzzeit-Leistungsspektren für die Charakterisierung von Wasserschallquellen. Konferenzband DAGA 2020, 603–606.

[2] Durbin, J., und Koopman, S.J.: Time Series Analysis by State Space Methods. 2. Ausgabe. Oxford University Press, Oxford, 2012.

[3] Shumway, R.H., und Stoffer, D.S.: An approach to time series smoothing and forecasting using the EM algorithm. Journal of Time Series Analysis, 3, 253–264, 1982.

[4] Nocedal, J., und Wright, S.: Numerical Optimization. 2. Ausgabe. Springer-Verlag, Berlin, 2006.

[5] Forst, C., Ginzkey, L. und Wille, P.: Audio-CD „Schall im Meer“. Forschungsanstalt der Bundeswehr für Wasserschall und Geophysik, Kiel, 2000.