

# Signalverarbeitung in der Physikalischen Geodäsie

Wolf-Dieter Schuh  
Institut für Geodäsie und Geoinformation, Universität Bonn  
Nussallee 17  
53115 Bonn  
schuh@uni-bonn.de

Author-created version of Schuh, W.-D. (2016)

**Zusammenfassung:** In diesem Kapitel werden Verfahren zur rechnerisch-analytischen Auswertung von Messreihen diskutiert. Dabei wird davon ausgegangen, dass die Messungen sowohl von deterministischen als auch von stochastischen Anteilen beeinflusst sind. Ziel einer geeigneten Modellierung ist die Trennung der einzelnen Anteile. Während der deterministische Anteil weitgehend durch geometrische und physikalische Zusammenhänge erklärt und parametrisiert werden kann, liegt über die Zusammensetzung und das Verhalten der stochastischen Einflüsse nur unpräzise Informationen vor. Diese Folge von Zufallsvariablen wird daher als stochastischer Prozess modelliert, dem Eigenschaften wie Stationarität, Homogenität und Isotropie zugeordnet werden können. Für den stochastischen Prozess werden unterschiedliche Darstellungsformen sowohl im Zeitbereich als auch im Frequenzbereich einander gegenübergestellt, wobei bei kovarianzstationären Prozessen auch die zweiten Momente mit einbezogen werden. Autokovarianzen im Zeitbereich und das Leistungsdichtespektrum im Frequenzbereich bilden somit zusammen mit dem Signal und dessen spektraler Darstellung ein Viereck - das Magische Quadrat. Dieser Name wurde gewählt, da vielfach geschlossene Formeln für die Umrechnung zwischen den vier Darstellungsformen gefunden werden können und somit der Weg für unterschiedlichste Modellierungs- und Berechnungsvarianten zur Berücksichtigung der Korrelationen im stochastischen Modell eröffnet werden. Als Optimierungsmodell wird das klassische Kollokationmodell herangezogen, wie in der Physikalischen Geodäsie üblich. Dieses Modell wird als bester linearer erwartungstreuer Prädiktor dargestellt, was eine direkte Gegenüberstellung zu den Kriging-Modellen (Simple-Kriging, Ordinary-Kriging und Universal-Kriging) aus der Geostatistik erlaubt. Für die numerische Implementierung der Korrelationen wird der Zugang über Kovarianzfunktionen und der Zugang über stochastische Prozesse näher diskutiert, die beide auch mit sehr großen Datenmengen noch effizient umgehen können. Zunächst wird auf Kovarianzfunktionen und deren mathematischen Eigenschaften näher eingegangen, der Zusammenhang zu Kovarianzmatrizen und deren positiver Definitheit hergestellt bevor dann spezielle Methoden zur Erstellung von finiten Kovarianzfunktionen erarbeitet werden. Finite Kovarianzfunktionen in  $\mathbb{R}$ ,  $\mathbb{R}^2$  und  $\mathbb{R}^3$  werden dargestellt und deren Anwendungen auf dem Kreis  $\mathbb{S}$  und der Kugel  $\mathbb{S}^2$  kurz diskutiert. Für regelmäßig abgetastete kovarianzstationäre Prozesse wird ein alternativer Zugang zur Beschreibung des stochastischen Modells über diskrete lineare Prozesse aufgezeigt. Durch das Magische Quadrat wird die Äquivalenz der beiden Zugänge bewiesen.

## Schlüsselwörter:

Stochastische Prozesse, Stationarität, finite Kovarianzfunktionen, Kollokation, Magisches Quadrat, AR-Prozess, Dekorrelationsfilter.

## 1 Einleitung

Neben der qualitativen Naturbeobachtung spielt die quantitative Erfassung der vielfältigen Prozesse im System Erde eine zentrale Rolle, denn **Messen schafft Wissen**. Durch Messungen lassen sich die Prozesse direkt oder indirekt beobachten und deren räumliche und/oder zeitliche Entwicklung analysieren. Bei der Durchführung von Messungen werden die zu quantifizierenden Parameter oft durch zusätzliche Signale überlagert, die sowohl Nutz- als auch Störsignale darstellen können. Das Beobachtungsmodell wird durch eine Kombination von deterministischen und stochastischen Anteilen beschrieben. In der Physikalischen Geodäsie wird vielfach die Methode der Kollokation nach Krarup (1969) und Moritz

(1972) verwendet und die Beobachtung  $\mathcal{L}_t$  durch

$$\mathcal{L}_t = \mathbf{A}_t \boldsymbol{\xi} + \mathcal{S}_t + \mathcal{N}_t \quad (1)$$

dargestellt. Dabei bezeichnen  $\mathcal{L}_t$  und  $\mathcal{S}_t$  Zufallsvariablen, die den Beobachtungsprozess bzw. den Signalanteil für eine bestimmte Situation  $t$  beschreiben.  $\mathbf{A}_t \in \mathbb{R}^{1 \times m}$  stellt den funktionalen Zusammenhang zwischen der Beobachtungsgröße  $\mathcal{L}_t$  und den  $m$  (wahren) Parametern  $\boldsymbol{\xi} \in \mathbb{R}^m$  her. Ein zumeist auftretendes rein zufälliges Störsignal (Rauschen) wird durch  $\mathcal{N}_t$  gekennzeichnet. Der Index  $t$  kann dabei sowohl einen bestimmten Ort als auch eine bestimmte Zeit oder eine sonstige Situationsbeschreibung repräsentieren. Die Menge aller gemessenen Situationen sammeln wir in  $T$  ( $t \in T$ ), wobei die Mächtigkeit  $\#T$  (=Anzahl der Situationen) mit  $n$  festgelegt ist,  $\#T = n$ . Die Sequenzen  $\{\mathcal{L}_t, t \in T\}$ ,  $\{\mathcal{S}_t, t \in T\}$  und  $\{\mathcal{N}_t, t \in T\}$  beschreiben somit stochastische Prozesse, deren Eigenschaften durch die (multivariate) Wahrscheinlichkeitsdichtefunktionen (PDF) bzw. deren Momente (z.B. Erwartungswert  $E\{\cdot\}$ , Varianz  $\Sigma\{\cdot\}$ , ...) beschrieben werden. Ordnet man die Beobachtungen, Signale und das Rauschen in einer geeigneten Reihenfolge und fasst sie in den Vektor  $\mathcal{L}$ ,  $\mathcal{S}$  und  $\mathcal{N}$  zusammen, so ergibt sich der multivariate Erwartungswert  $\mathbf{E}\{\cdot\}$  als vektorwertige Funktion und  $\Sigma\{\cdot\}$  kennzeichnet das multivariate zweite zentrale Moment, die Varianzen und Kovarianzen.

In Vektor-Matrix Schreibweise lässt sich das Kollokationsmodell durch

$$\mathcal{L} = \mathbf{A}\boldsymbol{\xi} + \mathcal{S} + \mathcal{N} \quad (2)$$

mit  $\mathcal{L}, \mathcal{S}, \mathcal{N} \in \mathbb{R}^n$ ;  $\boldsymbol{\xi} \in \mathbb{R}^m$ ;  $\mathbf{A} \in \mathbb{R}^{n \times m}$  zusammenfassen, wobei die Zufallsvektoren durch ihre ersten und zweiten Momente  $\mathbf{E}\{\mathcal{S}\}$ ,  $\Sigma\{\mathcal{S}\}$  bzw.  $\mathbf{E}\{\mathcal{N}\}$ ,  $\Sigma\{\mathcal{N}\}$  genauer beschrieben sind. Wir gehen für die weiteren Darstellungen davon aus, dass  $\mathcal{S}$  und  $\mathcal{N}$  unkorreliert sind,  $\Sigma\{\mathcal{S}, \mathcal{N}\} = \mathbf{0}_N$ .  $\mathbf{0}_N$  kennzeichnet eine quadratische Nullmatrix mit  $n$  Zeilen und Spalten. Somit ist die Varianz von  $\mathcal{L}$  als Summe der Einzelvarianzen  $\Sigma\{\mathcal{L}\} = \Sigma\{\mathcal{S}\} + \Sigma\{\mathcal{N}\}$  darstellbar. Die durchgeführten Messungen  $l$  werden nun als eine spezielle Realisierung der Zufallsvariablen  $\mathcal{L}$  betrachtet, womit auch  $s$  und  $n$  konkrete Realisierungen der Zufallsvektoren  $\mathcal{S}$  und  $\mathcal{N}$  bezeichnen,

$$l = \mathbf{A}\boldsymbol{\xi} + s + n. \quad (3)$$

In diesem Artikel wird konsequent zwischen Zufallsgrößen in kaligraphischen Schriftzeichen  $\mathcal{X}$ , wahren Werten in griechischen Zeichen  $\xi$  und Messungen oder Realisierungen der Zufallsvariablen in lateinischen (italic) Buchstaben  $x$  unterschieden. Vektoren, Matrizen und vektorwertige Funktionen werden durch Fettdruck  $\mathcal{X}$ ,  $\boldsymbol{\xi}$ ,  $x$ ,  $\mathbf{E}\{\mathcal{X}\}$  herausgehoben. Die Tilde (Schlange) über den Variablen weist immer auf ausgeglichene (optimale geschätzte) Größen hin. Die Varianz/Kovarianzen des Signals  $\Sigma\{\mathcal{S}\}$  und des Rauschens  $\Sigma\{\mathcal{N}\}$  für alle gemessenen Situationen  $t \in T$  mit  $\#T = n$  werden in den Matrizen  $\Sigma_{\mathcal{S}\mathcal{S}}$  bzw.  $\Sigma_{\mathcal{N}\mathcal{N}}$  zusammengefasst und die Kurzform  $\Sigma := \Sigma_{\mathcal{S}\mathcal{S}} + \Sigma_{\mathcal{N}\mathcal{N}}$  eingeführt.

Der Artikel ist wie folgt aufgebaut. Im Kap. 2 werden wir zunächst sehr allgemein auf stochastische Prozesse eingehen und uns dann speziell kovarianzstationären Prozessen widmen. Durch Anordnung mehrerer Darstellungsformen in einem einheitlichen Schema - *Magisches Quadrat* genannt - erstellen wir eine kompakte Übersicht über diese Darstellungsformen. Näheres Augenmerk wird im Kap. 2.1 auf regelmäßig abgetastete kovarianzstationäre Prozesse gelegt. Im Kap. 3 wird dann die Theorie zur Parameterschätzung (Kap. 3.1) und Signalschätzung (Kap. 3.2) durch Kollokation wiedergeben. Auf die Auswertung des Gesamtmodells werden wir in Kap. 3.3 eingehen und die Ergebnisse der Kollokationsmodelle den aus der Geostatistik weit verbreiteten Krigingmodellen gegenüberstellen. Im Kap. 4 werden unterschiedliche Zugänge zur Modellierung der Kovarianzen beschrieben, wobei einerseits auf den Zugang mit Kovarianzfunktionen (Kap. 4.1) bzw. finiten Kovarianzfunktionen (Kap. 4.2) und andererseits

die Modellierung über diskrete kovarianzstationäre Prozesse (Kap. 4.3) eingegangen wird, wo dann spezielle Dekorrelationsfilter (Kap. 4.4) zur effizienten Modellierung der Korrelationen für äquidistanten stochastische Prozesse entwickelt werden. Ein Fazit und Ausblick (Kap. 5) schließt dann diese Arbeit ab.

## 2 Stochastische Prozesse

Alle Messungen werden durch Signale überlagert, die sich aus vielen zufällig wirkenden Komponenten zusammensetzen. Die mathematische Beschreibung dieser zufälligen Vorgänge erfolgt über einen kontinuierlichen stochastischen Prozess  $\{\mathcal{S}(\mathbf{t}); \mathbf{t} \in \mathbb{R}^d\}$  oder wenn diskrete Abtastpunkte  $\mathbf{t}_i$  vorliegen durch einen diskreten stochastischen Prozess  $\{\mathcal{S}_{\mathbf{t}_i}; i \in \mathbb{Z}\}$ . Die Abtastpunkte  $\mathbf{t}_i$  beschreiben dabei einen Punkt im Raum und/oder in der Zeit, der durch  $d$ -Dimensionen aufgespannt ist und vielfach auf eine Region beschränkt ist.

Für jeden Abtastpunkt  $\mathbf{t}_i$  ist  $\mathcal{S}_{\mathbf{t}_i}$  eine Zufallsvariable, die durch die Randverteilung  $G_{\mathbf{t}_i}^{\mathcal{S}}(\mathbf{x}); \mathbf{x} \in \mathbb{R}$ , einer gemeinsamen (Wahrscheinlichkeits-)Verteilungsfunktion

$$F_{\{\mathbf{t}_i; i \in \mathbb{Z}\}}^{\mathcal{S}}(\mathbf{x}) = P\{\mathcal{S}_{\mathbf{t}_i} \leq \mathbf{x}; i \in \mathbb{Z}\} \quad (4)$$

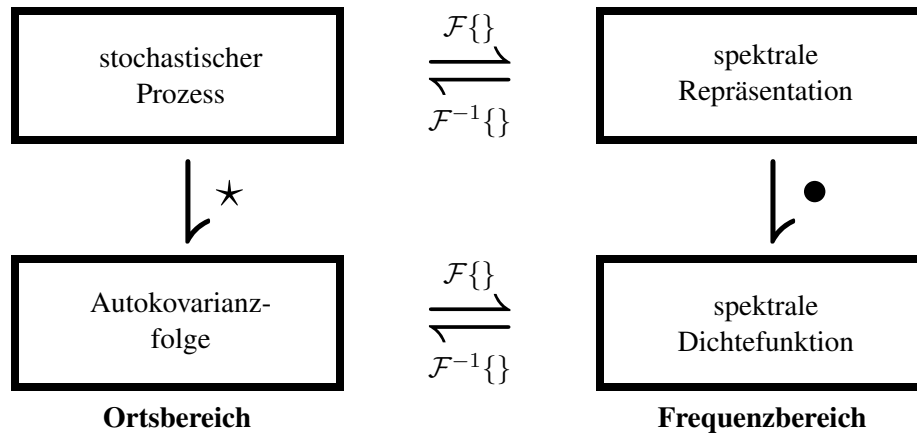
beschrieben wird. Die Variable  $x$  beschreibt die freie Variable der Randverteilungsfunktion  $G_{\mathbf{t}_i}^{\mathcal{S}}(x)$  für die Zufallsvariable  $\mathcal{S}_{\mathbf{t}_i}$ , wohingegen im Vektor  $\mathbf{x}$  die freien Variablen der gemeinsamen Verteilungsfunktion  $F_{\{\mathbf{t}_i; i \in \mathbb{Z}\}}^{\mathcal{S}}(\mathbf{x})$  zusammengefasst sind. Kovarianzen zwischen den Zufallsvariablen  $\mathcal{S}_{\mathbf{t}_j}$  und  $\mathcal{S}_{\mathbf{t}_k}$  für die Abtastpunkte  $\mathbf{t}_j$  und  $\mathbf{t}_k$  werden durch die gemeinsame Randverteilung  $G_{\mathbf{t}_j, \mathbf{t}_k}^{\mathcal{S}}(\mathbf{x})$  mit  $\mathbf{x} = [x_1 \ x_2]^T \in \mathbb{R}^2$  festgelegt, wobei  $x_1$  wieder die freie Variable in der Verteilungsfunktion für den ersten Abtastpunkt  $\mathbf{t}_i$  und  $x_2$  für  $\mathbf{t}_j$  kennzeichnet. Für höhere multivariate Momente gilt dies entsprechend.

Betrachten wir jedoch nur Prozesse, deren statistische Eigenschaften im Raum/Zeit unverändert bleiben, so beschränken wir uns auf *stationäre Prozesse*. Nach Khintchine (1934, S. 606) (siehe auch Brockwell und Davis (1991, Def 1.3.3)) ist ein stochastischer Prozess  $\{\mathcal{S}_{\mathbf{t}_i}; i \in \mathbb{Z}\}$  *stationär*, wenn die gemeinsame Randverteilungen bezüglich  $k$  aufeinanderfolgender Zufallsvariablen zu verschiedenen Zeitpunkten  $\mathbf{t}_i$  und  $\mathbf{t}_{i+h}$  ident sind,

$$G_{\mathbf{t}_i, \dots, \mathbf{t}_{i+k-1}}^{\mathcal{S}}(\mathbf{x}) = G_{\mathbf{t}_{i+h}, \dots, \mathbf{t}_{i+h+k-1}}^{\mathcal{S}}(\mathbf{x}), \quad k \in \mathbb{N}^+; i, h \in \mathbb{Z}; \mathbf{x} \in \mathbb{R}^k. \quad (5)$$

$h$  legt dabei einen beliebigen Verschiebeparameter fest. Für  $k = 1$  gilt, dass die Randverteilung  $G_{\mathbf{t}_i}^{\mathcal{S}}(x)$  für alle  $i$  unverändert bleibt und somit der Erwartungswert  $E\{\mathcal{S}_{\mathbf{t}_i}\}$ , die Varianz  $\Sigma\{\mathcal{S}_{\mathbf{t}_i}\}$  sowie alle höheren univariaten Momente  $E\{\mathcal{S}_{\mathbf{t}_i}^p\}$  der Zufallsvariablen  $\mathcal{S}_{\mathbf{t}_i}$  für alle Abtastpunkte  $\mathbf{t}_i$  unverändert bleiben. Für  $k = 2$  sind auch die bedingten Wahrscheinlichkeitsverteilungen zwischen jeweils zwei Abtastpunkten gleich. Somit bleiben auch die Kovarianzen zwischen jeweils zwei Abtastpunkten unverändert. Man spricht dann von *Stationarität der Ordnung 2*, *schwach stationären Prozessen* oder etwas präziser auch von *kovarianzstationären stochastischen Prozessen*. Für  $k = \infty$  spricht man von *stark stationären Prozessen*.

Bei geodätischen Beobachtungsprozessen nach (1) wird vielfach davon ausgegangen, dass der stochastische Prozess durch ein funktionales Modell  $\mathcal{A}\xi$  überlagert wird. Im Zusammenhang mit der Stationarität der Zeitreihe wird dieser deterministische Trend vielfach als Störgröße empfunden. Rummel (1975) spricht von einem *systematisch verfälschten Zufallsvorgang* und empfiehlt eine Trendabspaltung vorweg vorzunehmen, um die Stationarität zu gewährleisten. Nach Wold (1938) setzt sich ein kovarianzstationärer Prozess aus einem deterministischen Anteil und einem stochastischen Anteil (Innovationsprozess) zusammen, womit der Trendanteil in dieser Definition auch Bestandteil des stationären Prozesses ist. Beiden Argumentationen begegnet man oft und man sollte daher genau auf die Definitionen achten. Wir werden in weiterer Folge noch mehrfach auf diese Diskussion zurückkommen.



**Abb. 1:** Magisches Quadrat eines kovarianzstationären Prozesses

Betrachtet man *Gauß-Prozesse*, für welche die Verteilung aller Zufallsvariablen durch eine multivariate Normalverteilung festgelegt ist, haben *kovarianzstationäre Prozesse* eine besondere Bedeutung. Da die Randverteilungen einer multivariaten Normalverteilung wieder der Normalverteilung genügen, wird diese durch die beiden Parameter - Erwartungswert  $E\{\mathcal{S}_{t_i}\}$  und Autokovarianz  $\Sigma\{\mathcal{S}_{t_i}, \mathcal{S}_{t_j}\} = E\{(\mathcal{S}_{t_i} - E\{\mathcal{S}_{t_i}\})(\mathcal{S}_{t_j} - E\{\mathcal{S}_{t_j}\})\}$  - vollständig beschrieben. Um starke Stationarität für den *Gauß-Prozess* herzustellen genügen somit die Forderungen (Brockwell und Davis, 1991, Def 1.3.2)

$$\bullet E\{\mathcal{S}_{t_i}\} = \mu \quad \text{mit } \mu \dots \text{const.}, \quad (6)$$

$$\bullet E\{\mathcal{S}_{t_i}^2\} < \infty, \quad (7)$$

$$\bullet \Sigma\{\mathcal{S}_{t_i}\} = E\{(\mathcal{S}_{t_i} - E\{\mathcal{S}_{t_i}\})^2\} = \sigma^2 \quad \text{mit } \sigma^2 \dots \text{const. und} \quad (8)$$

$$\bullet \Sigma\{\mathcal{S}_{t_i}, \mathcal{S}_{t_j}\} = \Sigma\{\mathcal{S}_{t_{i+h}}, \mathcal{S}_{t_{j+h}}\} \quad \text{für } i, j, h \in \mathbb{Z}. \quad (9)$$

Zur Beschreibung der Eigenschaften eines kovarianzstationären stochastischen Prozesses werden oft sowohl Darstellungen im Ortsbereich als auch im Frequenzbereich genutzt. Neben dem *stochastischen Prozess* und der *Autokovarianzfolge* im Ortsbereich werden auch die *spektrale Repräsentation* des Prozesses und die *spektrale Verteilungs- oder Dichtefunktion* verwendet. Diese vier Darstellungsformen und ihre Beziehungen untereinander können sehr anschaulich in einer Figur zusammengefasst werden, die wir als "*Magisches Quadrat*" bezeichnen. In Abb. 1 bilden die vier Eckpunkte die vier unterschiedlichen Darstellungsformen des kovarianzstationären stochastischen Prozesses. Links ist der Ortsbereich und rechts der Frequenzbereich angeordnet. Die horizontalen Pfeile deuten den mathematischen Übergang vom Ortsbereich in den Frequenzbereich bzw. die Umkehrung an. Für diesen Übergang verwendet man oft den Begriff *Fouriertransformation*, was symbolisch mit  $\mathcal{F}\{\cdot\}$  angedeutet ist. Die Umkehroperation vom Frequenzbereich in den Ortsbereich wird als *inverse Fouriertransformation* bezeichnet und ist symbolisch mit  $\mathcal{F}^{-1}\{\cdot\}$  gekennzeichnet. Der Übergang von der oberen in die untere Zeile erfolgt im Ortsbereich durch einen Korrelationsoperator ( $\star$ ) und im Frequenzbereich gemäß dem Korrelationstheorem durch eine Multiplikation ( $\bullet$ ). Somit kann auf unterschiedlichen Wegen von einer Darstellungsform auf die andere übergangen werden, womit die Analogie zum *Magischen Quadrat* sehr nahe liegt.

Mit dem Magischen Quadrat können nun unterschiedliche Arten von Signalen erfasst werden. Die Signale können

- reell, komplex oder vektorwertig,
- kontinuierlich oder diskret, regelmäßig oder nicht regelmäßig,

- periodisch oder nicht periodisch,
- gefenstert oder unendlich ausgedehnt in Raum und/oder Zeit
- stochastisch oder deterministisch

sein, um hier nur einige Möglichkeiten aufzuzeigen. Davon unmittelbar abhängig muss eine entsprechende Form der *Fouriertransformation* in den Frequenzbereich definiert werden. Damit lassen sich die weiteren Darstellungsformen und Transformationen herleiten und somit das Magische Quadrat vervollständigen. Aber auch besondere Prozesse wie *autoregressive Prozesse* oder *Moving-Average Prozesse* können als Ausgangssignale vorgegeben werden und deren Darstellungsformen und Übergänge konkretisiert werden. Nicht in allen Fällen können geschlossene analytische Darstellungen für die Übergänge abgeleitet werden. Vielfach können auf unterschiedlichen Pfaden auch unterschiedliche Rechenformeln für ein und dieselbe Darstellungsform hergeleitet werden.

Bei automatisierten Messprozessen liegt zumeist eine regelmäßige Abtastung vor, sodass dem regelmäßig abgetasteten kovarianzstationären Prozess in vielen Anwendungen eine Sonderstellung zukommt. Daher werden wir diesen Prozess im Unterabschnitt 2.1 näher beleuchten.

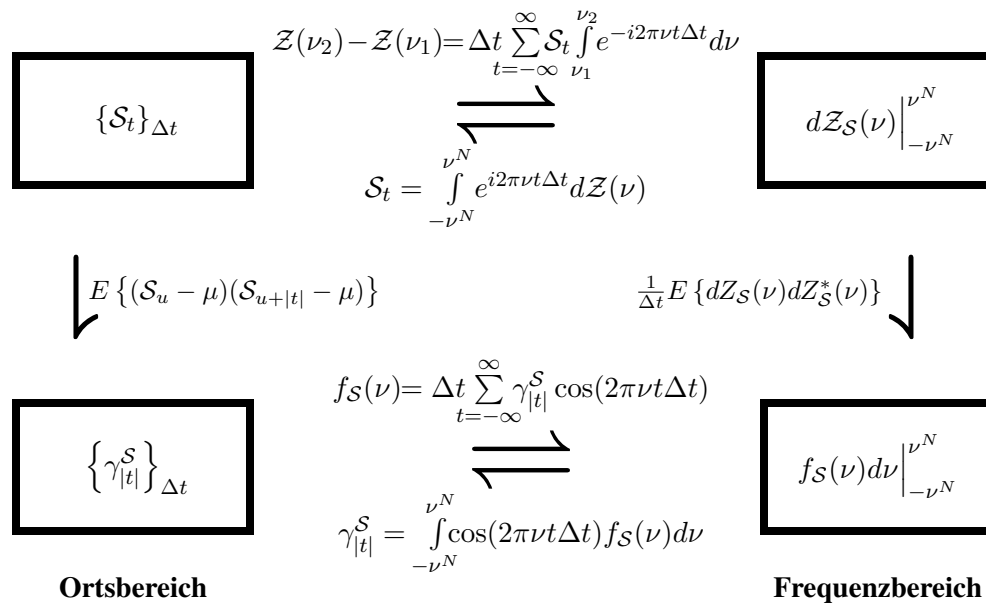
## 2.1 Regelmäßig abgetasteter kovarianzstationärer Prozess

Werden bei einem zeitdiskreten stochastischen Prozess  $\{\mathcal{S}_{t_i}; i \in \mathbb{Z}\}$  nur regelmäßige Zeitpunkte  $t_i = i\Delta t$  in Betracht gezogen, so beschreibt  $\{\mathcal{S}_{i\Delta t}; i \in \mathbb{Z}\}$  einen regelmäßig abgetasteten stochastischen Prozess, den wir im Folgenden kurz mit  $\{\mathcal{S}_t\}_{\Delta t}$  bezeichnen werden. Die Autokovarianzen  $\Sigma\{\mathcal{S}_{i\Delta t}, \mathcal{S}_{j\Delta t}\}$  sind wegen (9) nur vom Abstand  $|i-j|$  zwischen den Zeitpunkten abhängig. Die absoluten Zeitpunkte  $t_i = i\Delta t$  bzw.  $t_j = j\Delta t$  sind belanglos. Die *Autokovarianzfolge*

$$\gamma_{|t|}^{\mathcal{S}} = \gamma^{\mathcal{S}}(|i-j|\Delta t) = \Sigma\{\mathcal{S}_{t_i}, \mathcal{S}_{t_j}\} = E\{(\mathcal{S}_{t_i} - \mu)(\mathcal{S}_{t_j} - \mu)\} \quad (10)$$

wird kurz mit  $\gamma_{|t|}^{\mathcal{S}}$  bezeichnet, wobei  $\mathcal{S}$  den zugehörigen regelmäßig abgetasteten, stochastischen Prozess  $\{\mathcal{S}_t\}_{\Delta t}$  und  $t$  die Differenz  $t = i - j$  kennzeichnet. Der Absolutbetrag von  $t$  weist darauf hin, dass die Kovarianzfolge symmetrisch zum Nullpunkt ist, was aus der Stationaritätsforderung (9) unmittelbar folgt. Weiters sei darauf verwiesen, dass die Kovarianzfolge einer Folge von wahren (festen) Werten entspricht.

In Abb. 2 ist das Magische Quadrat für einen regelmäßig abgetasteten stochastischen Prozess dargestellt. Dem stochastischen Prozess und der Kovarianzfolge im Ortsbereich werden wieder die entsprechenden Darstellungsformen im Frequenzbereich gegenübergestellt. Die spektrale Repräsentation des regelmäßig abgetasteten stochastischen Prozesses im Ortsbereich, bildet ein kontinuierlicher, nur stückweise definierter, komplexwertiger stochastischer Prozess  $d\mathcal{Z}(\nu)$  im Frequenzbereich, welcher periodisch bezüglich des Intervalls  $[-\nu^N, \nu^N]$  ist (vgl. Krasbutter et al., 2015).  $\nu^N$  kennzeichnet die Nyquist Frequenz festgelegt durch  $\nu^N = \frac{1}{2\Delta t}$ . Eine *stochastische Fouriertransformation* beschreibt den Übergang vom Ortsbereich in den Frequenzbereich. Die Rücktransformation wird durch ein *stochastisches Fourierintegral* geleistet. Die vierte Darstellungsgröße bildet die spektrale Dichtefunktion (Spektraldichte)  $f_{\mathcal{S}}(\nu)$ , die reellwertig und periodisch bezüglich des Intervalls  $[-\nu^N, \nu^N]$  ist. Die diskrete Form des Wiener-Chintschin-Theorems (Buttkus, 2000, Kap. 8.2.2) definiert die Transformationen der Kovarianzfolge vom Ortsbereich in den Frequenzbereich. Da die Kovarianzfolge eine symmetrische Folge von reellen Zahlen darstellt, erfolgt der Übergang zur spektralen Dichtefunktion vereinfacht über eine *diskrete Kosinustransformation*. Daraus folgt, dass die spektrale Dichtefunktion ebenfalls eine gerade, kontinuierliche Funktion ist und die Rücktransformation durch eine *kontinuierliche Kosinustransformation* bewerkstelligt werden kann.



**Abb. 2:** Magisches Quadrat eines kovarianzstationären, regelmäßig abgetasteten stochastischen Prozesses. Oben links: stochastisches Signal, unten links: Autokovarianzfolge, oben rechts: spektrale Repräsentation des stochastischen Prozesses und unten rechts: spektrale Dichtefunktion (Spektraldichte). Die angeführten Formeln geben den Übergang in Pfeilrichtung wieder.

### 3 Kollokation

Bei der Verarbeitung von Zeitreihen oder raumbezogenen Daten zählt die Approximation, Interpolation und Extrapolation zu den Grundaufgaben. Unterschiedliche Begrifflichkeiten wie Prädiktion, Kollokation, Wiener-Kolmogorov-Filterung oder Kriging werden in den einzelnen Disziplinen verwendet und weisen auf die Pioniere dieses Forschungsgebietes Norbert Wiener (1894-1964), Andrej N. Kolmogorov (1903-1987), Danie H. Krige (1919-2013) hin. Während in der Geostatistik nach Matheron (1963) (siehe auch Cressie (1991), Wackernagel (2003)) vielfach von Kriging-Verfahren gesprochen wird, wird in der Signalverarbeitung eher auf die grundlegenden Arbeiten Kolmogorov (1941) und Wiener (1949) zurückgehend vom Wiener-Kolmogorov-Filter gesprochen. Das Kollokationsmodell wurde von Krarup (1969) und Moritz (1972) in die Geodäsie eingeführt und danach vielfach diskutiert (Grafarend, 1976; Rummel, 1976; Sansò, 1986; Schwarz, 1976; Wolf, 1974) und in Moritz (1980) ausführlich behandelt.

Aus der Sicht der Ausgleichsrechnung weisen diese unterschiedlichen Ansätze viele Gemeinsamkeiten aber auch einige Eigenheiten auf. Aus der deterministischen Sichtweise nach Wiener (1949, Kap. II) wird die Quadratsumme der Abweichungen für einen linearen Schätzer minimiert, während nach Kolmogorov (1941) die Varianz des Schätzfehlers durch einen bestangepassten Hilbertraum minimiert wird. Sowohl das Prinzip der *Methode der kleinsten Quadrate* als auch der BLUE (*best linear unbiased estimator - bester linearer erwartungstreuer Schätzer*) sind etabliert. Die Besonderheit der Anwendung bei stochastischen Prozessen ist jedoch dadurch gegeben, dass neben den Parametern des deterministischen Modells auch die Signalanteile bestmöglich geschätzt werden.

Sowohl bei den Kriging-Verfahren als auch bei der Kollokation unterscheidet man dabei unterschiedliche Zugänge. So wird beim *Simple-Kriging* analog zum (*Wiener-Kolmogorov*)-Prädiktion davon ausgegangen, dass das Signal einen Erwartungswert/Mittelwert Null über das gesamte Gebiet aufweist, während beim *Ordinary Kriging* ein konstanter, unbekannter Erwartungswert vorliegen kann. Ein nicht konstanter Erwartungswert wird durch eine Trendabspaltung mit deterministischen Modellen entweder im *Universal-Kriging* oder bei der

Kollokation berücksichtigt. In Abschnitt 3.3 und speziell in Tab. 3 werden wir nochmals auf diese Diskussion zurückkommen.

Charakteristisch für das Kriging Verfahren ist vielfach die Verwendung des Semivariogramms anstelle der Kovarianzfunktion. Während die Autokovarianzfunktion  $\gamma^S(t) = \Sigma \{ \mathcal{S}_{t_i}, \mathcal{S}_{t_j} \}$  über das zweite zentrale Moment (10) durch das Produkt der zentrierten Signale  $\mathcal{S}_{t_i}$  und  $\mathcal{S}_{t_j}$  definiert ist, werden die Korrelationen im Semivariogramm (halbe Varianz des Inkrementes)  $\bar{\gamma}^S(t) = \frac{1}{2} \Sigma \{ \mathcal{S}_{t_i} - \mathcal{S}_{t_j} \}$  durch Differenzen beschrieben. Der Parameter  $t$  legt dabei die Distanz zwischen den Aufpunkten  $\mathcal{S}_{t_i}$  und  $\mathcal{S}_{t_j}$  fest. Durch die Differenzbildung wird dabei ein oft vorliegender konstanter, unbekannter Trend (Mittelwert) in den Daten automatisch eliminiert. Werden die Kriging-Verfahren allerdings konsequent mit den Variogrammen durchgerechnet, so stößt man bei der Inversion der aus den Variogrammen gebildeten Korrelationsmatrix auf numerische Probleme, da die Diagonalglieder dieser symmetrischen Matrix oft sehr klein oder Null sind. Kovarianzmatrizen weisen hingegen eine zumeist dominante Hauptdiagonale auf, welche deren Berechnung/Inversion vereinfacht. Für die praktische Berechnung geht man daher vielfach vom Variogramm durch die Beziehung  $\bar{\gamma}^S(t) = \gamma^S(0) - \gamma^S(t)$  auf die Kovarianzen über und führt die numerische Berechnung mit Kovarianzmatrizen durch.

Als einen wesentlichen Unterschied zwischen Kollokation und Kriging bezeichnen Reguzzoni et al. (2005) die Tatsache, dass in der Kollokation auch beliebige Funktionale eines Signals in einem Guss mitverarbeitet werden können. Durch die Übertragung von Kovarianzfunktionen auf Funktionale des Signals (z.B. Ableitungen) durch ein verallgemeinertes Varianzfortpflanzungsgesetz (Moritz, 1973, S. 66) (Moritz, 1980, S. 171-173) kann dies bewerkstelligt werden (siehe auch Kap. 4.2). Sicher eine besondere Spezialität der Kollokationsverfahren. Allerdings werden in Reguzzoni et al. (2005) auch Wege aufgezeigt, wie die Erweiterung auf beliebige lineare Funktionale auch mit dem Variogramm-Ansatz umgesetzt werden kann.

Um dies kurz zusammenzufassen: Kollokationsverfahren erlauben sowohl eine Parameterschätzung als auch eine Schätzung/Prädiktion des Signals eines kovarianzstationären Prozesses, arbeiten mit Kovarianzfunktionen und erlauben eine konsistente Aufnahme von unterschiedlichen Funktionalen eines stochastischen Prozesses in den Schätzprozess. Vielfach wird das Argument vorgebracht, dass der Begriff Kollokation der Erdschwerfeldmodellierung vorbehalten ist. Viele weitere typische geodätische Anwendungen, wie zum Beispiel Oberflächenmodellierungen (z.B. Ozeantopographie,...) und die Modellierung von bewegten Objekten können jedoch auch mit Kollokationsmodellen ausgezeichnet bearbeitet werden. Somit sehe ich die Kollokation als ein allgemeines Werkzeug zur Lösung von geodätischen Ausgleichungsproblemen mit der speziellen Charakteristik stochastische Prozesse präzisieren zu können und habe daher diesen Namen als Kapitelüberschrift gewählt.

Moritz (1980, Kap. 16) zeigt zunächst zwei vollkommen unterschiedliche Zugänge zur Lösung des Kollokationsmodells (2) auf. Im Ansatz nach der *Kleinsten-Quadrate Kollokation* (LSC - Least-Squares Collocation) wählt Moritz einen algebraischen Zugang zur Lösung des linearen überbestimmten Gleichungssystem (3), welches sich nach der Durchführung von Messungen  $\ell$  als spezielle Realisierungen der Zufallsvariablen  $\mathcal{L}$  ergibt. Über die gewogene Quadratsumme der Signal- und Rauschanteile

$$\inf_{\tilde{\mathbf{x}}, \tilde{\mathbf{s}}, \tilde{\mathbf{n}}} (\mathbf{s}^T \Sigma_{SS}^{-1} \mathbf{s} + \mathbf{n}^T \Sigma_{NN}^{-1} \mathbf{n}) \quad (11)$$

wird das Infimum (Minimum) unter der Nebenbedingung

$$\mathbf{A}\mathbf{x} + \mathbf{s} + \mathbf{n} - \ell = \mathbf{0}_n \quad (12)$$

gebildet, wobei  $\mathbf{0}_n$  einen Nullvektor mit  $n$  Elementen kennzeichnet. Dies resultiert in den optimalen Parametern

$$\tilde{\mathbf{x}} = (\mathbf{A}^T \Sigma^{-1} \mathbf{A})^{-1} \mathbf{A}^T \Sigma^{-1} \ell \quad (13)$$

$$\tilde{\mathbf{s}} = \Sigma_{SS} \Sigma^{-1} (\ell - \mathbf{A}\tilde{\mathbf{x}}) \quad (14)$$

$$\tilde{\mathbf{n}} = \Sigma_{NN} \Sigma^{-1} (\ell - \mathbf{A}\tilde{\mathbf{x}}) . \quad (15)$$

Im zweiten Ansatz betrachtet Moritz das Kollokationsmodell (2) als statistisches Schätzproblem und bildet den *Besten Linearen Erwartungstreuen Prädiktor* (BLUP - Best Linear Unbiased Predictor). Dazu

wird der Schätzfehler  $\mathcal{E}_{\mathcal{X}}$  und  $\mathcal{E}_{\mathcal{S}}$  als Abweichung der geschätzten Parameter  $\mathcal{X}$  vom wahren Wert  $\xi$  bzw. der Signale  $\mathcal{S}$  vom Erwartungswert  $E\{\mathcal{S}\}$  definiert,

$$\mathcal{E}_{\mathcal{X}} := \mathcal{X} - \xi \quad \text{bzw.} \quad \mathcal{E}_{\mathcal{S}} := \mathcal{S} - E\{\mathcal{S}\}. \quad (16)$$

Der Schätzer sollte dabei *linear* von den Beobachtungen abhängen

$$\mathcal{X} = \mathbf{G}^T \mathcal{L} \quad \text{bzw.} \quad \mathcal{S} = \mathbf{H}^T \mathcal{L}. \quad (17)$$

Die Matrizen  $\mathbf{G} \in \mathbb{R}^{n \times m}$  und  $\mathbf{H} \in \mathbb{R}^{n \times n}$  beschreiben die Koeffizienten der Linearkombinationen mit denen die geschätzten Größen direkt aus den Messungen  $\mathcal{L}$  hergeleitet werden. Diese Koeffizienten können als Gewichte gesehen werden. Jede zu schätzende Größe wird also aus einem gewogenen Mittel direkt aus den Beobachtungen hergeleitet. Der  $j$ -te Spaltenvektor  $\mathbf{g}^{(j)}$  der Matrix  $\mathbf{G}$  (= Zeilenvektor in  $\mathbf{G}^T$ ) beschreibt die entsprechenden Gewichte zur Schätzung der Größe  $\mathcal{X}_j$ ,  $j = 1, \dots, m$ . Entsprechendes gilt für die Signale mit der  $i$ -ten Spalte  $\mathbf{h}^{(i)}$  aus  $\mathbf{H}$  für  $\mathcal{S}_i$ ,  $i = 1, \dots, n$ .

Der Schätzer sollte *erwartungstreu* sein, was durch das Verschwinden der Erwartungswerte der Schätzfehler

$$E\{\mathcal{E}_{\mathcal{X}}\} \stackrel{!}{=} \mathbf{0}_m \quad \text{bzw.} \quad E\{\mathcal{E}_{\mathcal{S}}\} \stackrel{!}{=} \mathbf{0}_n \quad (18)$$

gewährleistet wird.

Der *beste* Schätzer  $\tilde{\mathcal{X}}$  und  $\tilde{\mathcal{S}}$  für die Parameter und Signale wird über die minimale Varianzen  $\Sigma\{\mathcal{E}_{\tilde{\mathcal{X}}}\}$  und  $\Sigma\{\mathcal{E}_{\tilde{\mathcal{S}}}\}$  bzw. die entsprechenden Kovarianzmatrizen  $\Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}}$  und  $\Sigma_{\mathcal{E}_{\tilde{\mathcal{S}}}\mathcal{E}_{\tilde{\mathcal{S}}}}$  der Schätzfehler festgelegt. Die direkte Lösung dieser nichtskalaren Extremwertaufgabe erweist sich als schwierig, sodass der Optimalitätsbeweis indirekt geführt, indem gezeigt wird, dass

$$\mathbf{u}^T \left( \Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}} - \Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}} \right) \mathbf{u} \geq 0 \quad \text{bzw.} \quad \mathbf{v}^T \left( \Sigma_{\mathcal{E}_{\tilde{\mathcal{S}}}\mathcal{E}_{\tilde{\mathcal{S}}}} - \Sigma_{\mathcal{E}_{\tilde{\mathcal{S}}}\mathcal{E}_{\tilde{\mathcal{S}}}} \right) \mathbf{v} \geq 0 \quad (19)$$

für beliebige Vektoren  $\mathbf{u} \in \mathbb{R}^m$  und  $\mathbf{v} \in \mathbb{R}^n$  gilt, wobei  $\tilde{\mathcal{X}}$  und  $\tilde{\mathcal{S}}$  eine beliebige Lösung und  $\tilde{\mathcal{X}}$  und  $\tilde{\mathcal{S}}$  die optimale Lösung charakterisieren. Diese Forderung besagt, dass die Differenz der beiden Kovarianzmatrizen  $\Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}}$  und  $\Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}}$  positiv semidefinit sein muss (Moritz, 1980, S. 127-131). Um im Folgenden aber die Extremwertaufgabe direkt lösen zu können, führen wir ein Ersatzproblem ein und fordern, dass die Varianzen

$$\sigma_{\tilde{\mathcal{U}}}^2 = \mathbf{u}^T \Sigma_{\mathcal{E}_{\tilde{\mathcal{X}}}\mathcal{E}_{\tilde{\mathcal{X}}}} \mathbf{u} \quad \text{bzw.} \quad \sigma_{\tilde{\mathcal{V}}}^2 = \mathbf{v}^T \Sigma_{\mathcal{E}_{\tilde{\mathcal{S}}}\mathcal{E}_{\tilde{\mathcal{S}}}} \mathbf{v} \quad (20)$$

minimal werden sollen, wobei  $\tilde{\mathcal{U}} = \mathbf{u}^T \tilde{\mathcal{X}}$ ,  $\mathbf{u} \in \mathbb{R}^m$  bzw.  $\tilde{\mathcal{V}} = \mathbf{v}^T \tilde{\mathcal{S}}$ ,  $\mathbf{v} \in \mathbb{R}^n$  beliebige lineare Funktionale der optimalen Parameter  $\tilde{\mathcal{X}}$  und Signale  $\tilde{\mathcal{S}}$  sein können (Meissl, 1982, Sec. B.4.3). Diese Forderung scheint zunächst weit stärker zu sein, als die ursprüngliche Forderung nach den minimalen Kovarianzmatrizen. Wohingegen eine andere oft verwendete Forderung nach der Minimierung der Spur von  $\Sigma\{\mathcal{E}_{\tilde{\mathcal{X}}}\}$  und  $\Sigma\{\mathcal{E}_{\tilde{\mathcal{S}}}\}$  zunächst als schwächere Forderung eingeschätzt wird. Mathematisch führen jedoch beide skalaren Ersatzprobleme auf das idente Optimierungsproblem. Durch Einsetzen dieser Lösung kann dann gezeigt werden, dass diese Lösung auch optimal im Sinne der ursprünglich nichtskalaren Optimierungsaufgabe ist.

In Tabelle 1 werden die Forderungen für den *Besten Linearen Erwartungstreuen Prädiktor* (BLUP - Best Linear Unbiased Predictor) formelmäßig zusammengefasst, wobei  $\mathbf{H}$  und  $\mathbf{G}$  die freien Parameter der Optimierungsaufgabe repräsentieren.

### 3.1 Parameterschätzung im Kollokationsmodell

In diesem Abschnitt wollen wir auf die Parameterschätzung nach den Forderungen aus Tab. 1 eingehen. Wir benutzen dazu streng die Forderungen in der Reihenfolge: linear, erwartungstreu und best.



	Parameterschätzung	Signalschätzung
linear:	$\mathcal{X} = G^T \mathcal{L}$	$\mathcal{S} = H^T \mathcal{L}$
erwartungstreu:	$E \{ \mathcal{E}_{\mathcal{X}} \} \stackrel{!}{=} \mathbf{0}_m$	$E \{ \mathcal{E}_{\mathcal{S}} \} \stackrel{!}{=} \mathbf{0}_n$
best:	$\inf_{\tilde{G}} (u^T \Sigma_{\mathcal{E}_{\mathcal{X}}} u)$	$\inf_{\tilde{H}} (v^T \Sigma_{\mathcal{E}_{\mathcal{S}}} v)$
Nebenbedingung: $A\xi + \mathcal{S} + \mathcal{N} - \mathcal{L} = \mathbf{0}_n$		

**Table 1:** Zusammenfassung der Forderungen an das Kollokationsmodell als besten linearen erwartungstreuen Schätzer

**linear:** Nach Tab. 1 errechnen sich die geschätzten Parameter  $\mathcal{X}$  durch eine linearen Abbildung  $G^T$  direkt aus den Beobachtungen  $\mathcal{L}$ . Unter Beachtung der Nebenbedingung aus Tab. 1 bzw. (2) folgt somit der Zusammenhang

$$\mathcal{X} = G^T \mathcal{L} = G^T (A\xi + \mathcal{S} + \mathcal{N}) . \quad (21)$$

**erwartungstreu:** Der Erwartungswert  $E \{ \mathcal{E}_{\mathcal{X}} \}$  des Schätzfehlers der Parameter  $\mathcal{E}_{\mathcal{X}}$  (16) errechnet sich somit durch

$$\begin{aligned} E \{ \mathcal{E}_{\mathcal{X}} \} &= E \{ \mathcal{X} - \xi \} = \\ &= E \{ G^T (A\xi + \mathcal{S} + \mathcal{N}) - \xi \} = \\ &= (G^T A - \mathbb{1}_m) \xi + G^T E \{ \mathcal{S} \} + G^T E \{ \mathcal{N} \} , \end{aligned} \quad (22)$$

wobei  $\mathbb{1}_m$  eine Einheitsmatrix der Dimension  $m$  kennzeichnet. Die Forderung nach der Erwartungstreue des Schätzers bzw. Schätzfehlers

$$E \{ \mathcal{E}_{\mathcal{X}} \} \stackrel{!}{=} \mathbf{0}_m \quad \text{bzw.} \quad (G^T A - \mathbb{1}_m) \xi + G^T E \{ \mathcal{S} \} + G^T E \{ \mathcal{N} \} \stackrel{!}{=} \mathbf{0}_m \quad (23)$$

wird erreicht, wenn folgende Bedingungen

$$G^T A - \mathbb{1}_m \stackrel{!}{=} \mathbf{0}_m \quad (24)$$

$$G^T E \{ \mathcal{S} \} \stackrel{!}{=} \mathbf{0}_m \quad (25)$$

$$G^T E \{ \mathcal{N} \} \stackrel{!}{=} \mathbf{0}_m \quad (26)$$

erfüllt sind.  $\mathbf{0}_m$  kennzeichnet eine quadratische Matrix mit Nullelementen der Dimension  $m$ . Diese Bedingungen müssen also bei der Formulierung der Optimierungsaufgabe berücksichtigt werden. Setzt man allerdings wie üblich voraus, dass der Erwartungswert des Rauschens verschwindet  $E \{ \mathcal{N} \} = \mathbf{0}_n$  ist für die Einhaltung der Forderung (26) automatisch gesorgt. Andere Bedingungen müssen aber gegebenenfalls durch die Lagrange-Methode als Nebenbedingungen in die Optimierungsaufgabe eingeführt werden. Die Bedingung (24) besagt, dass  $G^T$  die Eigenschaften einer Linksinversen zur Matrix  $A$  erfüllen muss und ist als Nebenbedingung einzubringen. Für den Erwartungswert des Signals  $E \{ \mathcal{S} \}$  führt Moritz (1980, S. 100) den Begriff des *totalen Erwartungswerts* ein und setzt dann voraus, dass dieses Mittel über alle Erwartungswerte im Gebiet zu Null wird. Durch diese Annahme wird der Forderung (25) automatisch

genüge getan. Wir wollen hier ein anderes Konzept verfolgen und betrachten das Signal  $\mathcal{S}$  als stochastischen Prozess im Gebiet. Die Forderung nach einem stationären Verhalten im Gebiet (6) (siehe Abschnitt 2) führt dazu, dass der Erwartungswert des Signals unabhängig vom Ort immer den selben Wert besitzt, der im Folgenden mit  $E_{\mathcal{S}}$  gekennzeichnet wird. Somit gilt

$$\mathbf{E} \{ \mathcal{S} \} = \mathbf{1}_n E_{\mathcal{S}}, \quad (27)$$

wobei  $\mathbf{1}_n$  einen Einservektor mit  $n$  Elementen bezeichnet;  $\mathbf{1}_n := [1 \dots, 1]^T$ ,  $\mathbf{1}_n \in \mathbb{R}^n$ . Die Forderung (25) kann nun in zwei Bedingungen aufgespalten werden

$$\mathbf{G}^T \mathbf{1}_n \stackrel{!}{=} \mathbf{0}_m \quad \text{oder} \quad E_{\mathcal{S}} \stackrel{!}{=} 0, \quad (28)$$

wobei die Einhaltung einer Bedingung ausreicht um die Erwartungstreue zu gewährleisten. Ist der Erwartungswert  $E_{\mathcal{S}}$  gleich Null, dann kann wie bei Moritz (1980, Kap. 16) verfahren werden. Wenn dies nicht vorausgesetzt werden kann, ist die alternative Bedingung von (28) als Nebenbedingung in die Optimierungsaufgabe aufzunehmen. Diese alternative Bedingung besagt, dass die Summe jeder Zeile in  $\mathbf{G}^T$  gleich Null sein muss, damit der konstante Erwartungswert des Signals ausgeblendet wird. Somit wird eine Fallunterscheidung für die folgenden Betrachtungen erforderlich.

Speziell bei modernen Sensoren (z.B. bei den Beschleunigungsmessern von GOCE) kann nicht immer gewährleistet werden, dass der Erwartungswert des Messrauschens/signals verschwindet und somit die Voraussetzung  $E_{\mathcal{S}} = 0$  immer erfüllt ist. In diesen Fällen kann die alternative Forderung  $\mathbf{G}^T \mathbf{1}_n \stackrel{!}{=} \mathbf{0}_m$  von (28) eingesetzt werden, damit die Parameterschätzung erwartungstreu bleibt. Die hier beschriebene Vorgehensweise bei  $E_{\mathcal{S}} \neq 0$  ist äquivalent mit erweiterten funktionalen Modellen, wo ein konstanter Anteil im Signal über einen Zusatzparameter erfasst wird. Man vergleiche auch den Zugang über stochastische Impulse Jäggi (2007) oder kurze Bahnbögen Mayer-Gürr (2008), wo Zusatzparameter epochenweise eingesetzt werden um die Variabilität des Erwartungswerts des Signals zu modellieren.

**best:** Zur Berechnung des *besten* Schätzers benötigen wir zunächst die Varianz des Schätzfehlers  $\Sigma \{ \mathcal{E}_{\mathcal{X}} \}$ , die sich berechnet durch

$$\begin{aligned} \Sigma \{ \mathcal{E}_{\mathcal{X}} \} &= \mathbf{E} \{ \mathcal{E}_{\mathcal{X}} \mathcal{E}_{\mathcal{X}}^T \} = \\ &= \mathbf{E} \{ (\mathcal{X} - \xi)(\mathcal{X} - \xi)^T \} = \\ &= \mathbf{E} \{ (\mathbf{G}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) - \xi)(\mathbf{G}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) - \xi)^T \} = \\ &= \mathbf{E} \left\{ ((\mathbf{G}^T \mathbf{A} - \mathbf{1}_m)\xi + \mathbf{G}^T (\mathcal{S} + \mathcal{N})) ((\mathbf{G}^T \mathbf{A} - \mathbf{1}_m)\xi + \mathbf{G}^T (\mathcal{S} + \mathcal{N}))^T \right\}. \end{aligned} \quad (29)$$

Da Erwartungstreue für den Schätzer gefordert wird, müssen die Bedingungen (24), (25) und (26) eingehalten werden. Durch Einsatz von (24) vereinfacht sich (29) zu

$$\Sigma \{ \mathcal{E}_{\mathcal{X}} \} = \mathbf{G}^T \mathbf{E} \{ (\mathcal{S} + \mathcal{N})(\mathcal{S} + \mathcal{N})^T \} \mathbf{G}. \quad (30)$$

Ersetzt man nun das zweite Moment von Signal und Rauschen  $\mathbf{E} \{ (\mathcal{S} + \mathcal{N})(\mathcal{S} + \mathcal{N})^T \}$  durch das zweite zentrale Moment  $\Sigma \{ \mathcal{S} + \mathcal{N} \}$  so errechnet sich unter Beachtung von  $\mathbf{E} \{ \mathcal{N} \} = \mathbf{0}$  und  $\mathbf{E} \{ \mathcal{S} \} = \mathbf{1}_n E_{\mathcal{S}}$  der Zusammenhang

$$\mathbf{E} \{ (\mathcal{S} + \mathcal{N})(\mathcal{S} + \mathcal{N})^T \} = \Sigma \{ \mathcal{S} + \mathcal{N} \} + \mathbf{1}_n \mathbf{1}_n^T E_{\mathcal{S}}^2. \quad (31)$$

Somit kann (30) umgeschrieben werden zu

$$\Sigma \{ \mathcal{E}_{\mathcal{X}} \} = \mathbf{G}^T \Sigma \{ \mathcal{S} + \mathcal{N} \} \mathbf{G} + \mathbf{G}^T \mathbf{1}_n \mathbf{1}_n^T \mathbf{G} E_{\mathcal{S}}^2. \quad (32)$$

Der zweite Term wird wegen Bedingung (25) bzw. (28) zu Null, unabhängig welche der Forderungen angesetzt wird, denn entweder ist  $E_{\mathcal{S}} = 0$  und damit auch  $E_{\mathcal{S}}^2$ , oder es gilt  $\mathbf{G}^T \mathbf{1}_n = \mathbf{0}_m$  und damit auch  $\mathbf{G}^T \mathbf{1}_n \mathbf{1}_n^T \mathbf{G} = \mathbf{0}_m$ . Die Varianz des Schätzfehlers der Parameter  $\Sigma \{\mathcal{E}_{\mathcal{X}}\}$  für einen erwartungstreuen Schätzer  $\mathcal{X}$  lässt sich somit durch

$$\Sigma \{\mathcal{E}_{\mathcal{X}}\} = \mathbf{G}^T \Sigma \{\mathcal{S} + \mathcal{N}\} \mathbf{G} = \mathbf{G}^T \Sigma \mathbf{G} \quad (33)$$

ausdrücken, wobei wir wieder  $\Sigma := \Sigma \{\mathcal{S} + \mathcal{N}\}$  verwenden.

**Bester Schätzer  $\tilde{\mathcal{X}}$  bei  $E_{\mathcal{S}} \stackrel{!}{=} 0$ :** Zur Ermittlung des besten Schätzers ist nun die quadratische Form

$$\mathbf{u}^T \Sigma \{\mathcal{E}_{\mathcal{X}}\} \mathbf{u} = \mathbf{u}^T \mathbf{G}^T \Sigma \mathbf{G} \mathbf{u} = \sum_{t=1}^m u_t^2 \mathbf{g}^{(t)T} \Sigma \mathbf{g}^{(t)} \quad (34)$$

mit Nebenbedingung (24) zu minimieren. Die Bedingungen (25) und (26) werden automatisch durch die getroffenen Annahmen  $E_{\mathcal{S}} \stackrel{!}{=} 0$  bzw.  $\mathbf{E} \{\mathcal{N}\} \stackrel{!}{=} \mathbf{0}_n$  erfüllt. Die Nebenbedingung (24) schreiben wir dabei in transponierter Form und verwenden wieder die Summenschreibweise

$$\mathbf{A}^T \mathbf{G} - \mathbf{1}_m = \mathbf{0}_m \quad \text{bzw.} \quad \mathbf{A}^T \mathbf{g}^{(t)} - \mathbf{e}^{(t)} = \mathbf{0}_m, \quad t = 1, \dots, m. \quad (35)$$

$\mathbf{e}^{(t)}$  kennzeichnet den  $t$ -ten Spaltenvektor der Einheitsmatrix  $\mathbf{1}_m$ . Nach der Lagrange-Methode setzt sich die zu minimierende Funktion aus der Minimumsbedingung (34) und den Nebenbedingungen (35) zusammen

$$\phi \left( \mathbf{g}^{(t)}, \mathbf{k}^{(t)} \right) = \sum_{t=1}^m u_t^2 \mathbf{g}^{(t)T} \Sigma \mathbf{g}^{(t)} + 2 \sum_{t=1}^m u_t^2 \mathbf{k}^{(t)T} \left( \mathbf{A}^T \mathbf{g}^{(t)} - \mathbf{e}^{(t)} \right), \quad (36)$$

wobei  $\mathbf{K} = [\mathbf{k}^{(1)}, \dots, \mathbf{k}^{(m)}]$  die Lagrange-Multiplikatoren kennzeichnen. Die Gradienten der Funktion  $\phi$  errechnen sich mit

$$\frac{1}{2} \nabla_{\mathbf{g}^{(t)}} \phi = u_t^2 \left( \Sigma \mathbf{g}^{(t)} + \mathbf{A} \mathbf{k}^{(t)} \right), \quad t = 1, \dots, m \quad (37)$$

$$\frac{1}{2} \nabla_{\mathbf{k}^{(t)}} \phi = u_t^2 \left( \mathbf{A}^T \mathbf{g}^{(t)} - \mathbf{e}^{(t)} \right), \quad t = 1, \dots, m \quad (38)$$

$$(39)$$

und werden stationär durch die Bedingungen

$$\left. \begin{array}{l} \Sigma \tilde{\mathbf{g}}^{(t)} + \mathbf{A} \tilde{\mathbf{k}}^{(t)} \stackrel{!}{=} \mathbf{0}_n, \quad t = 1, \dots, m \\ \mathbf{A}^T \tilde{\mathbf{g}}^{(t)} - \mathbf{e}^{(t)} \stackrel{!}{=} \mathbf{0}_m, \quad t = 1, \dots, m \end{array} \right\} \left\{ \begin{array}{l} \Sigma \tilde{\mathbf{G}} + \mathbf{A} \tilde{\mathbf{K}} \stackrel{!}{=} \mathbf{0}_{n \times m} \\ \mathbf{A}^T \tilde{\mathbf{G}} - \mathbf{1}_m \stackrel{!}{=} \mathbf{0}_m. \end{array} \right. \quad (40)$$

da  $u_t$  beliebige Werte annehmen kann. Aus den in (40) auf der rechten Seite in Matrixschreibweise zusammengefassten Bedingungen können direkt die Normalgleichungen

$$\begin{bmatrix} \Sigma & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0}_m \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}} \\ \tilde{\mathbf{K}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n \times m} \\ \mathbf{1}_m \end{bmatrix}. \quad (41)$$

aufgestellt werden. Die numerische Lösung der Normalgleichungen kann vielfach sehr effizient berechnet werden, wenn dünnbesetzte oder Toeplitz-Systeme vorliegen. Sofern  $\Sigma$  regulär ist<sup>1</sup>, kann das Gleichungssystem (40) auch analytisch gelöst werden. Aus der ersten Gleichung folgt

$$\tilde{\mathbf{G}} = -\Sigma^{-1} \mathbf{A} \tilde{\mathbf{K}}. \quad (42)$$

<sup>1</sup>Für die eindeutige Lösung des verallgemeinerten Dreieckssystems (41) besteht lediglich die Forderung, dass der Rang des erweiterten Systems  $[\Sigma | \mathbf{A}]$  und auch der Rang von  $\mathbf{A}$  voll sein muss,  $\text{Rang}(\Sigma | \mathbf{A}) = n$  und  $\text{Rang}(\mathbf{A}) = m$ .

Eingesetzt in die zweite Gleichung in (40) errechnen sich die Lagrange-Multiplikatoren mit

$$\tilde{\mathbf{K}} = - (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \quad (43)$$

woraus aus (42)

$$\tilde{\mathbf{G}} = \boldsymbol{\Sigma}^{-1} \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \quad (44)$$

folgt. Der bester erwartungstreue Schätzer unter Annahme von  $E\{\mathcal{N}\} = \mathbf{0}$  und  $E_{\mathcal{S}} = 0$  errechnet sich somit aus

$$\tilde{\boldsymbol{\chi}} = \tilde{\mathbf{G}}^T \boldsymbol{\mathcal{L}} = (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}}. \quad (45)$$

Die Varianz des Schätzfehlers  $\boldsymbol{\Sigma}\{\boldsymbol{\mathcal{E}}_{\tilde{\boldsymbol{\chi}}}\}$  folgt aus (33) mit

$$\boldsymbol{\Sigma}\{\boldsymbol{\mathcal{E}}_{\tilde{\boldsymbol{\chi}}}\} = \tilde{\mathbf{G}}^T \boldsymbol{\Sigma} \tilde{\mathbf{G}}. \quad (46)$$

Durch Einsetzen von (44) und Vereinfachung erhält man

$$\boldsymbol{\Sigma}\{\boldsymbol{\mathcal{E}}_{\tilde{\boldsymbol{\chi}}}\} = (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1}. \quad (47)$$

Das funktionale Modell  $\mathbf{A}\tilde{\boldsymbol{\chi}}$  erklärt den parametrischen Anteil des stochastischen Prozesses. Die restlichen Anteile

$$\widetilde{\Delta\boldsymbol{\mathcal{L}}} = \boldsymbol{\mathcal{L}} - \mathbf{A}\tilde{\boldsymbol{\chi}} \quad (48)$$

bezeichnet man als *verkürzte Beobachtungen* oder auch *Residuen*. Die verkürzten Beobachtungen können auch direkt von  $\boldsymbol{\mathcal{L}}$  abgeleitet werden, indem man (45) in (48) einsetzt

$$\widetilde{\Delta\boldsymbol{\mathcal{L}}} = \left( \mathbb{1}_n - \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \right) \boldsymbol{\mathcal{L}} = \left( \boldsymbol{\Sigma} - \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}}. \quad (49)$$

Nach Meissl (1982, Kap. A.5) verwenden wir Projektoren um diese Beziehung kompakt auszudrücken und auch geometrisch besser interpretierbar zu machen. Dazu führen wir die Projektoren

$$\boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} := \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \quad (50)$$

und

$$\boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} := \boldsymbol{\Sigma} - \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T = \boldsymbol{\Sigma} - \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \quad (51)$$

ein, die eine Projektion in den Spaltenraum  $S(\mathbf{A})$  bzw. Spaltenorthogonalraum  $S^\perp(\mathbf{A})$  der Matrix  $\mathbf{A}$  unter Beachtung der Metrik  $\boldsymbol{\Sigma}^{-1}$  gewährleisten. Man beachte, dass durch die Einführung des verallgemeinerten Skalarproduktes

$$\langle \mathbf{A}, \mathbf{B} \rangle_M := \mathbf{A}^T \mathbf{M} \mathbf{B} \quad (52)$$

für beliebige kompatible Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  mit einer positiv semidefiniten, symmetrischen Metrik  $\mathbf{M}$  spezielle Eigenschaften wie *Symmetrie*, *Idempotenz*, *Orthogonalität* und *Vollständigkeit* für diese Projektoren gelten (siehe Tab. 2).

Durch die Projektorenschreibweise können nun die Anteile des funktionalen Modells  $\mathbf{A}\tilde{\mathcal{X}}$  und der verkürzten Beobachtungen  $\widetilde{\Delta\mathcal{L}}$  sehr einfach geometrisch interpretiert werden. Unter Beachtung von (45) errechnet sich

$$\mathbf{A}\tilde{\mathcal{X}} = \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathcal{L} = \left\langle \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \mathcal{L} \right\rangle_{\boldsymbol{\Sigma}^{-1}} \quad (53)$$

und die verkürzten Beobachtungen  $\widetilde{\Delta\mathcal{L}}$  nach (49) mit

$$\widetilde{\Delta\mathcal{L}} = \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}^{-1} \mathcal{L} = \left\langle \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \mathcal{L} \right\rangle_{\boldsymbol{\Sigma}^{-1}}. \quad (54)$$

Somit zeigt sich, dass der funktionale Anteil  $\mathbf{A}\tilde{\mathcal{X}}$  als jener Anteil interpretiert werden kann, der entsteht, wenn der Beobachtungsvektor  $\mathcal{L}$  in den Spaltenraum der Matrix  $\mathbf{A}$  unter Beachtung der Metrik  $\boldsymbol{\Sigma}^{-1}$  projiziert wird. Für die verkürzten Beobachtungen werden durch den Projektor  $\boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$  alle Anteile von  $\mathcal{L}$  ausgeblendet, die sich im Spaltenraum  $S(\mathbf{A})$  befinden, sodass nur noch die Anteile des Spaltenorthogonalraums  $S^\perp(\mathbf{A})$  enthalten bleiben.

	$\boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} := \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T$	$\boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} := \boldsymbol{\Sigma} - \mathbf{A} (\mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^T$
Symmetrie	$\boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} = \left( \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right)^T$	$\boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} = \left( \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right)^T$
Idempotenz	$\left\langle \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right\rangle_{\boldsymbol{\Sigma}^{-1}} = \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$	$\left\langle \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right\rangle_{\boldsymbol{\Sigma}^{-1}} = \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$
Orthogonalität	$\left\langle \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right\rangle_{\boldsymbol{\Sigma}^{-1}} = \mathbf{0}_n$	$\left\langle \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \right\rangle_{\boldsymbol{\Sigma}^{-1}} = \mathbf{0}_n$
Vollständigkeit	$\boldsymbol{\Sigma} = \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} + \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$	
	$\langle \mathbf{A}, \mathbf{B} \rangle_M := \mathbf{A}^T \mathbf{M} \mathbf{B}$	

**Tabelle 2:** Eigenschaften der Projektoren in den Spaltenraum  $\boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$  bzw. Spaltenorthogonalraum  $\boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}$  der Matrix  $\mathbf{A}$  zur Metrik  $\boldsymbol{\Sigma}^{-1}$  und Definition des verallgemeinerten Skalarproduktes  $\langle \mathbf{A}, \mathbf{B} \rangle_M$  zur symmetrischen, positiv semidefiniten Metrik  $\mathbf{M}$ , für  $\mathbf{M} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{A} \in \mathbb{R}^{n \times m}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times p}$ .

Die Varianz des funktionalen Anteils  $\mathbf{A}\tilde{\mathcal{X}}$  und der verkürzten Beobachtungen  $\widetilde{\Delta\mathcal{L}}$  folgen aus dem Varianzfortpflanzungsgesetz. Unter Beachtung der Idempotenz der Projektoren bezüglich der Metrik  $\boldsymbol{\Sigma}^{-1}$  folgt sofort

$$\boldsymbol{\Sigma} \left\{ \mathbf{A}\tilde{\mathcal{X}} \right\} = \boldsymbol{\Pi}_{S(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}} \quad \text{bzw.} \quad \boldsymbol{\Sigma} \left\{ \widetilde{\Delta\mathcal{L}} \right\} = \boldsymbol{\Pi}_{S^\perp(\mathbf{A})}^{\boldsymbol{\Sigma}^{-1}}. \quad (55)$$

Diese Ergebnisse entsprechen den Resultaten von Moritz (1980, Kap. 17), wobei unterschiedliche Herleitungen verwendet werden. Der hier verwendete Weg über eine Extremwertaufgabe erlaubt die Erweiterung des Systems auch auf den Fall, wo die Voraussetzung  $E_{\mathcal{S}} \stackrel{!}{=} 0$  nicht erfüllt ist (siehe Kommentar auf Seite 10) und die Forderung  $\mathbf{G}^T \mathbf{1}_n = \mathbf{0}_m$  aus (28) zusätzlich in die Extremwertaufgabe einzubringen ist.

**Bester Schätzer für  $\tilde{\mathcal{X}}$  bei  $E_{\mathcal{S}} \neq 0$ :** Da wir hier nicht davon ausgehen wollen, dass der Erwartungswert des Signals  $E_{\mathcal{S}}$  gleich Null ist, muss zur Erhaltung der Erwartungstreue des Schätzers für  $\tilde{\mathcal{X}}$  die Bedingung  $\mathbf{G}^T \mathbf{1}_n = \mathbf{0}_m$  aus (28) als weitere Nebenbedingung in die Extremwertaufgabe aufgenommen werden. Es wird wieder die Varianz (34) eines beliebigen linearen Funktionals  $\mathcal{U} = \mathbf{u}^T \tilde{\mathcal{X}}$  minimiert. Neben der Bedingung (35) wird somit zusätzlich die Bedingung

$$\mathbf{G}^T \mathbf{1}_n = \mathbf{0}_m \quad \text{bzw.} \quad \mathbf{1}_n^T \mathbf{g}^{(t)} = 0, \quad t = 1, \dots, m \quad (56)$$

eingebraucht. Nach der Lagrange-Methode ergibt sich die zu minimierende Funktion mit

$$\phi(\mathbf{g}^{(t)}, \mathbf{k}^{(t)}, l_t) = \sum_{t=1}^m u_t^2 \mathbf{g}^{(t)T} \Sigma \mathbf{g}^{(t)} + 2 \sum_{t=1}^m u_t^2 \mathbf{k}^{(t)T} (\mathbf{A}^T \mathbf{g}^{(t)} - \mathbf{e}^{(t)}) + 2 \sum_{t=1}^m u_t^2 l_t \mathbf{1}_n^T \mathbf{g}^{(t)}. \quad (57)$$

wobei  $\mathbf{l} \in \mathbb{R}^m$  einen Vektor mit Lagrange-Multiplikatoren symbolisiert. Die Gradienten bzw. partielle Ableitungen errechnen sich mit

$$\frac{1}{2} \nabla_{\mathbf{g}^{(t)}} \phi = u_t^2 (\Sigma \mathbf{g}^{(t)} + \mathbf{A} \mathbf{k}^{(t)} + \mathbf{1}_n l_t), \quad t = 1, \dots, m \quad (58)$$

$$\frac{1}{2} \nabla_{\mathbf{k}^{(t)}} \phi = u_t^2 (\mathbf{A}^T \mathbf{g}^{(t)} - \mathbf{e}^{(t)}), \quad t = 1, \dots, m \quad (59)$$

$$\frac{1}{2} \frac{\partial \phi}{\partial l_t} = u_t^2 \mathbf{1}_n^T \mathbf{g}^{(t)}, \quad t = 1, \dots, m \quad (60)$$

und werden stationär durch die Bedingungen

$$\left. \begin{array}{l} \Sigma \tilde{\mathbf{g}}^{(t)} + \mathbf{A} \tilde{\mathbf{k}}^{(t)} + \mathbf{1}_n \tilde{l}_t \stackrel{!}{=} \mathbf{0}_n, \quad t = 1, \dots, m \\ \mathbf{A}^T \tilde{\mathbf{g}}^{(t)} - \mathbf{e}^{(t)} \stackrel{!}{=} \mathbf{0}_m, \quad t = 1, \dots, m \\ \mathbf{1}_n^T \tilde{\mathbf{g}}^{(t)} \stackrel{!}{=} 0, \quad t = 1, \dots, m \end{array} \right\} \left\{ \begin{array}{l} \Sigma \tilde{\mathbf{G}} + \mathbf{A} \tilde{\mathbf{K}} + \mathbf{1} \tilde{\mathbf{l}}^T \stackrel{!}{=} \mathbf{0}_{n \times m} \\ \mathbf{A}^T \tilde{\mathbf{G}} - \mathbf{1}_m \stackrel{!}{=} \mathbf{0}_m \\ \mathbf{1}^T \tilde{\mathbf{G}} \stackrel{!}{=} \mathbf{0}_m^T \end{array} \right. \quad (61)$$

dargestellt, wobei die links angeführten Einzelbedingungen wieder in Matrix-Vektor-Form zusammengefasst sind. Die Lösung kann implizit über die Normalgleichungen

$$\begin{bmatrix} \Sigma & \mathbf{A} & \mathbf{1}_n \\ \mathbf{A}^T & \mathbf{0}_m & \mathbf{0}_m \\ \mathbf{1}_n^T & \mathbf{0}_m^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}} \\ \tilde{\mathbf{K}} \\ \tilde{\mathbf{l}}^T \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n \times m} \\ \mathbf{1}_m \\ \mathbf{0}_m^T \end{bmatrix} \quad (62)$$

gefunden werden oder explizit über die Lösung der Gleichungen

$$\tilde{\mathbf{G}} = \Sigma^{-1} [\mathbf{A} \quad \mathbf{1}_n] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m^T \end{bmatrix} \quad (63)$$

berechnet werden, wenn  $\Sigma^{-1}$  existiert. Die beste erwartungstreue lineare Schätzung für  $E_{\mathcal{S}} \neq 0$  errechnet sich somit durch

$$\tilde{\mathcal{X}} = \tilde{\mathbf{G}}^T \mathcal{L} = [\mathbf{1}_m \quad \mathbf{0}_m] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} \Sigma^{-1} \mathcal{L}. \quad (64)$$

Zur Lösung der Inversen der blockweise gegebenen Normalgleichungen kann die Identität

$$\begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix} \begin{bmatrix} \mathbf{N}_{11}^{(-1)} & \mathbf{N}_{12}^{(-1)} \\ \mathbf{N}_{21}^{(-1)} & \mathbf{N}_{22}^{(-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_m & \mathbf{0}_n \\ \mathbf{0}_n^T & 1 \end{bmatrix} \quad (65)$$

genutzt werden, wobei  $N_{ii}^{(-1)}$  die Blöcke in der inversen Matrix kennzeichnen. Durch Ausmultiplikation und Auflösung erhält man für die Elemente in der ersten Zeile

$$N_{11}^{(-1)} = \left( \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{A} - \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n (\mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{A} \right)^{-1} \quad (66)$$

$$N_{21}^{(-1)} = -N_{11}^{(-1)} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n (\mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n)^{-1}. \quad (67)$$

Die geschätzten Parameter  $\tilde{\boldsymbol{\mathcal{X}}}$  aus (64) können somit durch

$$\tilde{\boldsymbol{\mathcal{X}}} = \left( N_{11}^{(-1)} \mathbf{A}^T - N_{21}^{(-1)} \mathbf{1}_n^T \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}} \quad (68)$$

berechnet werden. Unter Beachtung von (67) erhält man

$$\tilde{\boldsymbol{\mathcal{X}}} = N_{11}^{(-1)} \left( \mathbf{A}^T - \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n (\mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}_n^T \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}} \quad (69)$$

beziehungsweise

$$\tilde{\boldsymbol{\mathcal{X}}} = N_{11}^{(-1)} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \left( \boldsymbol{\Sigma} - \mathbf{1}_n (\mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}_n^T \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}} \quad (70)$$

Durch Einführung der Projektoren

$$\boldsymbol{\Pi}_{S(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} := \mathbf{1}_n (\mathbf{1}_n^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}_n^T \quad \text{bzw.} \quad \boldsymbol{\Pi}_{S^\perp(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} = \boldsymbol{\Sigma} - \boldsymbol{\Pi}_{S(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} \quad (71)$$

kann  $N_{11}^{(-1)}$  definiert durch (66) umgeschrieben werden zu

$$N_{11}^{(-1)} = \left( \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Pi}_{S^\perp(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}^{-1} \mathbf{A} \right)^{-1}. \quad (72)$$

Somit kann (70) umgeformt werden zu

$$\tilde{\boldsymbol{\mathcal{X}}} = \left( \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Pi}_{S^\perp(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Pi}_{S^\perp(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mathcal{L}} \quad (73)$$

und man erkennt, dass durch den Übergang auf die neue Metrik

$$\bar{\boldsymbol{\Sigma}}^{-1} := \boldsymbol{\Sigma}^{-1} \boldsymbol{\Pi}_{S^\perp(\mathbf{1}_n)}^{\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}^{-1} \quad (74)$$

die Schätzung (64) für  $\tilde{\boldsymbol{\mathcal{X}}}$  wieder die gewohnte Form

$$\tilde{\boldsymbol{\mathcal{X}}} = \left( \mathbf{A}^T \bar{\boldsymbol{\Sigma}}^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^T \bar{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\mathcal{L}} \quad (75)$$

annimmt. Die in (74) definierte Metrik  $\bar{\boldsymbol{\Sigma}}^{-1}$  ist somit unmittelbar interpretierbar als die Metrik, die dafür sorgt, dass der konstante Signalanteil bei der Parameterschätzung berücksichtigt wird. Aus geometrischer Sicht wird der Parameterraum  $S(\mathbf{A})$  um den Spaltenvektor  $\mathbf{1}_n$  verkleinert, sodass die Schätzung (75) invariant gegenüber einer konstanten Verschiebung  $E_S$  wird, da diese Verschiebung ausgeblendet wird.

Die Varianz des Schätzfehlers  $\boldsymbol{\Sigma} \{ \boldsymbol{\mathcal{E}}_{\tilde{\boldsymbol{\mathcal{X}}}} \}$  folgt aus (75) unmittelbar durch Varianzfortpflanzung. Unter Beachtung der Idempotenz der Projektoren (siehe Tab. 2) folgt

$$\boldsymbol{\Sigma} \{ \boldsymbol{\mathcal{E}}_{\tilde{\boldsymbol{\mathcal{X}}}} \} = \left( \mathbf{A}^T \bar{\boldsymbol{\Sigma}}^{-1} \mathbf{A} \right)^{-1}. \quad (76)$$

Die Inverse von  $\mathbf{A}^T \bar{\Sigma}^{-1} \mathbf{A}$  existiert, wenn der Vektor  $\mathbf{1}_n$  nicht im Spaltenraum von  $\mathbf{A}$  liegt,  $\mathbf{1}_n \notin S(\mathbf{A})$ . Dieser Vektor würde durch die Metrik aus den Normalgleichungen ausgeblendet, was einen Randdefekt zu Folge hätte. Diese Aussage kann auch direkt aus der verallgemeinerten Dreiecksgestalt der Normalgleichungsmatrix (62) abgelesen werden. Eine eindeutige Lösung existiert nur, wenn

$$\text{Rang} \begin{bmatrix} \Sigma \\ \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} = n \quad \text{und} \quad \text{Rang} [\mathbf{A} \mid \mathbf{1}] = m + 1 \quad (77)$$

gewährleistet ist. Daraus ist ersichtlich, dass eine eindeutige Lösung nur dann existiert, wenn  $\mathbf{1}_n$  nicht im Spaltenraum der Matrix  $\mathbf{A}$  liegt,  $\mathbf{1}_n \notin S(\mathbf{A})$  und  $\mathbf{A}$  vollen Spaltenrang aufweist,  $\text{rang}(\mathbf{A}) = m$ . Wird im deterministischen Modell bereits ein konstanter Term mitgeschätzt, so kann dieser konstante Term nicht vom Signalanteil getrennt werden, was zu singulären Normalgleichungen und einer Lösungsmannigfaltigkeit führt. Es gilt zu beachten, dass trotz der kompakten Schätzformel (75) für  $\tilde{\mathcal{X}}$  es vielfach numerisch günstiger sein kann die Schätzung aus dem Gesamtsystem (62) vorzunehmen, da Toeplitz und dünnbesetzte Strukturen effizient genutzt werden können.

### 3.2 Signalschätzung im Kollokationsmodell

Zur Ermittlung des besten linearen erwartungstreuen Schätzer für das Signal  $\mathcal{S}$  benutzen wir wieder dasselbe Vorgehen wie bei der Parameterschätzung. Wir bilden zunächst einen linearen Schätzer

$$\mathcal{S} = \mathbf{H}^T \mathcal{L} = \mathbf{H}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) . \quad (78)$$

Der Erwartungswert des Schätzfehlers des Signals  $\mathcal{E}_{\mathcal{S}} = \mathcal{S} - E\{\mathcal{S}\}$  nach (16) errechnet sich durch

$$\begin{aligned} E\{\mathcal{E}_{\mathcal{S}}\} &= E\{\mathcal{S} - E\{\mathcal{S}\}\} = \\ &= E\{\mathbf{H}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) - E\{\mathcal{S}\}\} = \\ &= \mathbf{H}^T \mathbf{A}\xi + (\mathbf{H}^T - \mathbf{1}_n) E\{\mathcal{S}\} + \mathbf{H}^T E\{\mathcal{N}\} . \end{aligned} \quad (79)$$

Die Forderung nach der Erwartungstreue des Schätzfehlers  $\mathcal{E}_{\mathcal{S}}$

$$E\{\mathcal{E}_{\mathcal{S}}\} \stackrel{!}{=} \mathbf{0}_n \quad \text{bzw.} \quad \mathbf{H}^T \mathbf{A}\xi + (\mathbf{H}^T - \mathbf{1}_n) E\{\mathcal{S}\} + \mathbf{H}^T E\{\mathcal{N}\} \stackrel{!}{=} \mathbf{0}_n \quad (80)$$

wird durch folgende Bedingungen

$$\mathbf{H}^T \mathbf{A} \stackrel{!}{=} \mathbb{O}_{n \times m} \quad (81)$$

$$(\mathbf{H}^T - \mathbf{1}_n) E\{\mathcal{S}\} \stackrel{!}{=} \mathbf{0}_n \quad (82)$$

$$\mathbf{H}^T E\{\mathcal{N}\} \stackrel{!}{=} \mathbf{0}_n \quad (83)$$

erreicht, wobei Bedingung (82) unter Beachtung der Stationarität des Signals (27) in zwei alternativ mögliche Bedingungen

$$\mathbf{H}^T \mathbf{1}_n - \mathbf{1}_n \stackrel{!}{=} \mathbf{0}_n \quad \text{oder} \quad E_{\mathcal{S}} \stackrel{!}{=} 0 \quad (84)$$

zerlegt wird und Bedingung (83) durch die Forderung  $E\{\mathcal{N}\} \stackrel{!}{=} \mathbf{0}$  zu erfüllen ist. Die Varianz des Schätzfehlers  $\Sigma\{\mathcal{E}_{\mathcal{S}}\}$  errechnet sich aus

$$\begin{aligned} \Sigma\{\mathcal{E}_{\mathcal{S}}\} &= E\{\mathcal{E}_{\mathcal{S}} \mathcal{E}_{\mathcal{S}}^T\} = \\ &= E\{(\mathbf{H}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) - E\{\mathcal{S}\})(\mathbf{H}^T (\mathbf{A}\xi + \mathcal{S} + \mathcal{N}) - E\{\mathcal{S}\})^T\} . \end{aligned} \quad (85)$$



Beachtet man (81) und löst den Erwartungswertoperator auf, so erhält man

$$\begin{aligned} \Sigma \{ \mathcal{E}_{\mathcal{S}} \} &= \mathbf{H}^T \mathbf{E} \{ (\mathcal{S} + \mathcal{N})(\mathcal{S} + \mathcal{N})^T \} \mathbf{H} - \mathbf{H}^T \mathbf{E} \{ (\mathcal{S} + \mathcal{N})^T \mathcal{S}^T \} - \\ &\quad - \mathbf{E} \{ \mathcal{S}(\mathcal{S} + \mathcal{N})^T \} \mathbf{H} + \mathbf{E} \{ \mathcal{S} \mathcal{S}^T \} . \end{aligned} \quad (86)$$

Ersetzt man die zweiten Momente durch die zweiten zentralen Momente

$$\mathbf{E} \{ (\mathcal{S} + \mathcal{N})(\mathcal{S} + \mathcal{N})^T \} = \Sigma \{ \mathcal{S} + \mathcal{N} \} + \mathbf{1}_n \mathbf{1}_n^T E_{\mathcal{S}}^2 \quad (87)$$

$$\mathbf{E} \{ (\mathcal{S} + \mathcal{N}) \mathcal{S}^T \} = \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} + \mathbf{1}_n \mathbf{1}_n^T E_{\mathcal{S}}^2 \quad (88)$$

$$\mathbf{E} \{ \mathcal{S}(\mathcal{S} + \mathcal{N})^T \} = \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} + \mathbf{1}_n \mathbf{1}_n^T E_{\mathcal{S}}^2 \quad (89)$$

$$\mathbf{E} \{ \mathcal{S} \mathcal{S}^T \} = \Sigma \{ \mathcal{S} \} + \mathbf{1}_n \mathbf{1}_n^T E_{\mathcal{S}}^2 \quad (90)$$

so erhält man

$$\Sigma \{ \mathcal{E}_{\mathcal{S}} \} = \mathbf{H}^T \Sigma \{ \mathcal{S} + \mathcal{N} \} \mathbf{H} - \mathbf{H}^T \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} - \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \mathbf{H} + \Sigma \{ \mathcal{S} \} , \quad (91)$$

wobei  $\Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}$  die Kovarianz zwischen  $\mathcal{S} + \mathcal{N}$  und  $\mathcal{S}$  darstellt und  $\Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \}$  die transponierte Form bildet. Um das Kriterium *best* zu erfüllen muss die Varianz  $\Sigma \{ \mathcal{E}_{\mathcal{S}} \}$  minimal werden, wobei die Fälle  $E_{\mathcal{S}} = 0$  und  $E_{\mathcal{S}} \neq 0$  nach den alternativ möglichen Bedingungen (84) zu unterscheiden sind.

**Bester Schätzer  $\tilde{\mathcal{S}}$  bei  $E_{\mathcal{S}} \stackrel{!}{=} 0$ :** Um den besten linearen erwartungstreuen Schätzer  $\tilde{\mathcal{S}}$  unter der Annahme  $\mathbf{E} \{ \mathcal{N} \} = \mathbf{0}$  und  $E_{\mathcal{S}} = 0$  zu ermitteln, ist die Varianz  $\Sigma \{ \mathbf{v}^T \Sigma \{ \mathcal{E}_{\mathcal{S}} \} \mathbf{v} \}$  der Linearkombination  $\mathbf{v}^T \tilde{\mathcal{S}}$  für beliebige  $\mathbf{v}$  zu minimieren. Somit ist die quadratische Form

$$\begin{aligned} \mathbf{v}^T \Sigma \{ \mathcal{E}_{\mathcal{S}} \} \mathbf{v} &= \mathbf{v}^T (\mathbf{H}^T \Sigma \{ \mathcal{S} + \mathcal{N} \} \mathbf{H} - \mathbf{H}^T \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} - \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \mathbf{H} + \Sigma \{ \mathcal{S} \}) \mathbf{v} = \\ &= \sum_{t=1}^n v_t^2 \mathbf{h}^{(t)T} \Sigma \mathbf{h}^{(t)} - 2 \sum_{t=1}^n v_t^2 \Sigma \{ \mathcal{S}_t, \mathcal{S} + \mathcal{N} \} \mathbf{h}^{(t)} + \sum_{t=1}^n v_t^2 \Sigma \{ \mathcal{S}_t \} , \quad t = 1, \dots, n \end{aligned} \quad (92)$$

unter der Nebenbedingung (81) stationär zu machen, wobei hier mit  $\mathcal{S}_t$  wieder auf die einzelne Komponente des Zufallsvektors  $\mathcal{S}$  übergegangen wird. Die Kovarianzmatrix  $\Sigma \{ \mathcal{S}_t, \mathcal{S} + \mathcal{N} \}$  reduziert sich auf einen liegenden Vektor mit den Kovarianzen zwischen der Zufallsvariablen  $\mathcal{S}_t$  und dem Zufallsvektor  $\mathcal{S} + \mathcal{N}$ . Die Nebenbedingung (81) schreiben wir dabei in transponierter Form und verwenden für die Darstellung die einzelnen Spaltenvektoren  $\mathbf{h}^{(t)}$  der Matrix  $\mathbf{H}$

$$\mathbf{A}^T \mathbf{H} = \mathbf{0}_{m \times n} \quad \text{bzw.} \quad \mathbf{A}^T \mathbf{h}^{(t)} = \mathbf{0}_m, \quad t = 1, \dots, n . \quad (93)$$

Nach der Lagrange-Methode ergibt sich die zu minimierende Funktion

$$\begin{aligned} \phi \left( \mathbf{h}^{(t)}, \mathbf{m}^{(t)} \right) &= \sum_{t=1}^n v_t^2 \mathbf{h}^{(t)T} \Sigma \mathbf{h}^{(t)} - 2 \sum_{t=1}^n v_t^2 \Sigma \{ \mathcal{S}_t, \mathcal{S} + \mathcal{N} \} \mathbf{h}^{(t)} + \\ &\quad + \sum_{t=1}^n v_t^2 \Sigma \{ \mathcal{S}_t \} + 2 \sum_{t=1}^n v_t^2 \mathbf{m}^{(t)T} \mathbf{A}^T \mathbf{h}^{(t)} , \end{aligned} \quad (94)$$

wobei  $\mathbf{M} = [\mathbf{m}^{(1)}, \dots, \mathbf{m}^{(n)}]$ ,  $\mathbf{M} \in \mathbb{R}^{m \times n}$  die Lagrange-Multiplikatoren kennzeichnen. Die Gradienten

$$\frac{1}{2} \nabla_{\mathbf{h}^{(t)}} \phi = u_t^2 \left( \Sigma \mathbf{h}^{(t)} - \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S}_t \} \mathbf{h}^{(t)} + \mathbf{A} \mathbf{m}^{(t)} \right) , \quad t = 1, \dots, n \quad (95)$$

$$\frac{1}{2} \nabla_{\mathbf{m}^{(t)}} \phi = u_t^2 \left( \mathbf{A}^T \mathbf{h}^{(t)} \right) , \quad t = 1, \dots, n \quad (96)$$

werden stationär durch die Forderungen

$$\left. \begin{array}{l} \Sigma \tilde{\mathbf{h}}^{(t)} + \Sigma \{ \mathbf{S} + \mathcal{N}, \mathcal{S}_t \} + \mathbf{A} \tilde{\mathbf{m}}^{(t)} \stackrel{!}{=} \mathbf{0}_n, \quad t=1, \dots, n \\ \mathbf{A}^T \tilde{\mathbf{h}}^{(t)} \stackrel{!}{=} \mathbf{0}_m, \quad t=1, \dots, n \end{array} \right\} \left\{ \begin{array}{l} \Sigma \tilde{\mathbf{H}} - \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \} + \mathbf{A} \tilde{\mathbf{M}} \stackrel{!}{=} \mathbf{0}_n \\ \mathbf{A}^T \tilde{\mathbf{H}} \stackrel{!}{=} \mathbf{0}_{m \times n} \end{array} \right. \quad (97)$$

Somit ergeben sich die Normalgleichungen zu

$$\begin{bmatrix} \Sigma & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0}_m \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{H}} \\ \tilde{\mathbf{M}} \end{bmatrix} = \begin{bmatrix} \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \} \\ \mathbf{0}_{m \times n} \end{bmatrix}. \quad (98)$$

Diese können numerisch nach  $\tilde{\mathbf{H}}$  aufgelöst werden oder bei regulärer Matrix  $\Sigma$  analytisch weiter bearbeitet werden. Durch Umformungen der Normalgleichungen (98) errechnet sich

$$\tilde{\mathbf{H}} = \Sigma^{-1} \underbrace{(\Sigma - \mathbf{A}(\mathbf{A}^T \Sigma^{-1} \mathbf{A})^{-1} \mathbf{A}^T)}_{\Pi_{S^\perp(A)}^{\Sigma^{-1}}} \Sigma^{-1} \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \} \quad (99)$$

und der beste Schätzer  $\tilde{\mathbf{S}}$  ist festgelegt durch

$$\tilde{\mathbf{S}} = \tilde{\mathbf{H}}^T \mathcal{L} = \Sigma \{ \mathbf{S}, \mathbf{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \Sigma^{-1} \mathcal{L}. \quad (100)$$

Unter Verwendung der verkürzten Beobachtungen  $\widetilde{\Delta \mathcal{L}}$  (54) führt dies zur kompakten Darstellung

$$\tilde{\mathbf{S}} = \Sigma \{ \mathbf{S}, \mathbf{S} + \mathcal{N} \} \Sigma^{-1} \widetilde{\Delta \mathcal{L}}. \quad (101)$$

In der Schreibweise mit verallgemeinerten Skalarprodukten (52) erhält man

$$\tilde{\mathbf{S}} = \left\langle \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \}, \widetilde{\Delta \mathcal{L}} \right\rangle_{\Sigma^{-1}} \quad \text{mit} \quad \widetilde{\Delta \mathcal{L}} = \left\langle \Pi_{S^\perp(A)}^{\Sigma^{-1}}, \tilde{\mathcal{L}} \right\rangle_{\Sigma^{-1}}, \quad (102)$$

und erkennt, dass sich der Schätzer  $\tilde{\mathbf{S}}$  einfach durch die Korrelation (Skalarprodukt) des Vektors der Kovarianzen (Kovarianzfunktion) mit den verkürzten Beobachtungen unter der Metrik  $\Sigma^{-1}$  ergibt. Die Varianz des Schätzers  $\Sigma \{ \tilde{\mathbf{S}} \}$  errechnet sich über Varianzfortpflanzung  $\Sigma \{ \tilde{\mathbf{S}} \} = \tilde{\mathbf{H}}^T \Sigma \{ \mathcal{L} \} \tilde{\mathbf{H}}$  unter Beachtung der Idempotenz von  $\Pi_{S^\perp(A)}^{\Sigma^{-1}}$  mit

$$\Sigma \{ \tilde{\mathbf{S}} \} = \Sigma \{ \mathbf{S}, \mathbf{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \Sigma^{-1} \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \} \quad (103)$$

beziehungsweise ausgedrückt als verallgemeinertes Skalarprodukt unter Einführung der alternativen Metrik (74)

$$\Sigma \{ \tilde{\mathbf{S}} \} = \left\langle \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \}, \Sigma \{ \mathbf{S} + \mathcal{N}, \mathbf{S} \} \right\rangle_{\bar{\Sigma}^{-1}} \quad \text{mit} \quad \bar{\Sigma}^{-1} = \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \Sigma^{-1}. \quad (104)$$

Die Varianz des Schätzfehler  $\Sigma \{ \mathcal{E}_{\tilde{\mathbf{S}}} \}$  kann über (91) durch Einsetzen von (99) berechnet werden. Nach entsprechenden Vereinfachungen erreicht man

$$\Sigma \{ \mathcal{E}_{\tilde{\mathbf{S}}} \} = \Sigma \{ \mathbf{S} \} - \Sigma \{ \tilde{\mathbf{S}} \}. \quad (105)$$

**Beste Schätzer  $\tilde{\mathcal{S}}$  bei  $E_{\mathcal{S}} \neq 0$ :** Um den besten Schätzer  $\tilde{\mathcal{S}}$  bei  $E_{\mathcal{S}} \neq 0$  zu ermitteln muss wieder die quadratische Funktion (92) minimiert werden, wobei als zusätzliche Nebenbedingung (93) berücksichtigt werden muss. Die zu minimierende Funktion (94) erweitert sich also um diese Nebenbedingung und lautet

$$\begin{aligned} \phi(\mathbf{h}^{(t)}, \mathbf{m}^{(t)}, n_t) &= \sum_{t=1}^n v_t^2 \mathbf{h}^{(t)T} \Sigma \mathbf{h}^{(t)} - 2 \sum_{t=1}^n v_t^2 \Sigma \{\mathcal{S}_t, \mathcal{S} + \mathcal{N}\} \mathbf{h}^{(t)} + \\ &+ \sum_{t=1}^n v_t^2 \Sigma \{\mathcal{S}_t\} + 2 \sum_{t=1}^n v_t^2 \mathbf{m}^{(t)T} \mathbf{A}^T \mathbf{h}^{(t)} + 2 \sum_{t=1}^n v_t^2 n_t \left( \mathbf{1}_n^T \mathbf{h}^{(t)} - 1 \right), \end{aligned} \quad (106)$$

wobei  $\mathbf{n} = [n_1, n_2, \dots, n_n]^T$ ,  $\mathbf{n} \in \mathbb{R}^n$  zusätzliche Lagrange-Multiplikatoren darstellen. Setzt man die Gradienten  $\frac{1}{2} \nabla_{\mathbf{h}^{(t)}} \phi$  und  $\frac{1}{2} \nabla_{\mathbf{m}^{(t)}} \phi$  bzw. die partiellen Ableitungen  $\frac{\partial \phi}{\partial n_t}$  dieser Funktion wieder zu Null und ordnet die  $t$  Bedingungen in Matrizen an, so ergeben sich die Normalgleichungen mit

$$\begin{bmatrix} \Sigma & \mathbf{A} & \mathbf{1}_n \\ \mathbf{A}^T & \mathbf{O}_m & \mathbf{0}_m \\ \mathbf{1}^T & \mathbf{0}_m^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{H}} \\ \tilde{\mathbf{M}} \\ \tilde{\mathbf{n}}^T \end{bmatrix} = \begin{bmatrix} \Sigma \{\mathcal{S} + \mathcal{N}, \mathcal{S}\} \\ \mathbf{O}_{m \times n} \\ \mathbf{1}_n^T \end{bmatrix}. \quad (107)$$

Löst man die Normalgleichungen nach den Parametern  $\tilde{\mathbf{H}}$  auf, errechnet sich die Matrix  $\tilde{\mathbf{H}}$  für den besten linearen erwartungstreuen Schätzer mit

$$\begin{aligned} \tilde{\mathbf{H}} &= \Sigma^{-1} \left( \Sigma - [\mathbf{A} \ \mathbf{1}] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} \right) \Sigma^{-1} \Sigma \{\mathcal{S} + \mathcal{N}, \mathcal{S}\} + \\ &+ \Sigma^{-1} [\mathbf{A} \ \mathbf{1}] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{O}_{M \times N} \\ \mathbf{1}_n^T \end{bmatrix}. \end{aligned} \quad (108)$$

Durch Einführung der Projektoren für den um den Vektor  $\mathbf{1}_n$  erweiterten Parameterraum  $S(\mathbf{A})$

$$\Pi_{S(\mathbf{A}|\mathbf{1}_n)}^{\Sigma^{-1}} := [\mathbf{A} \ \mathbf{1}_n] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} \quad (109)$$

$$\Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)}^{\Sigma^{-1}} := \Sigma - \Pi_{S(\mathbf{A}|\mathbf{1}_n)}^{\Sigma^{-1}} \quad (110)$$

ergibt sich  $\tilde{\mathcal{S}} = \tilde{\mathbf{H}}^T \mathcal{L}$  zunächst zu

$$\begin{aligned} \tilde{\mathcal{S}} &= \Sigma \{\mathcal{S}, \mathcal{S} + \mathcal{N}\} \Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)}^{\Sigma^{-1}} \Sigma^{-1} \mathcal{L} + \\ &+ [\mathbf{O}_{n \times m} \ \mathbf{1}_n] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} \Sigma^{-1} \mathcal{L}. \end{aligned} \quad (111)$$

Der zweite Term von (111) stellt die ausgeglichene Schätzung des Erwartungswertes des Signal multipliziert mit einem Vektor  $\mathbf{1}_n$  dar

$$\mathbf{1}_n \tilde{\mathcal{X}}_{E_{\mathcal{S}}} := [\mathbf{O}_{n \times m} \ \mathbf{1}_n] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} \Sigma^{-1} \mathcal{L}. \quad (112)$$

Durch Verwendung der Inversenidentität analog zu (65) kann (112) weiter aufgelöst werden. Somit kann der Ausdruck

$$[\mathbf{O}_{n \times m} \ \mathbf{1}_n] \begin{bmatrix} \mathbf{A}^T \Sigma^{-1} \mathbf{A} & \mathbf{A}^T \Sigma^{-1} \mathbf{1}_n \\ \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} & \mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{1}_n^T \end{bmatrix} = \mathbf{1}_n \mathbf{N}_{21}^{(-1)} \mathbf{A}^T + \mathbf{1}_n \mathbf{N}_{22}^{(-1)} \mathbf{1}_n^T \quad (113)$$

substituiert werden durch die entsprechenden Blöcke  $N_{21}^{(-1)}$  bzw.  $N_{22}^{(-1)}$  in der Inversen. Durch Lösung von (65) nach der zweiten Zeile erhält man

$$N_{22}^{(-1)} = \left( \mathbf{1}_n^T \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \mathbf{1}_n \right)^{-1} \quad (114)$$

$$N_{21}^{(-1)} = -N_{22}^{(-1)} \mathbf{1}_n^T \Sigma^{-1} \mathbf{A} (\mathbf{A}^T \Sigma^{-1} \mathbf{A})^{-1} . \quad (115)$$

Damit kann (113) umgeformt werden zu

$$\mathbf{1}_n \left( \mathbf{1}_n^T \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \mathbf{1}_n \right)^{-1} \mathbf{1}_n^T \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \mathcal{L} , \quad (116)$$

sodass (112) vereinfacht geschrieben werden kann

$$\mathbf{1}_n \tilde{\mathcal{X}}_{ES} = \mathbf{1}_n \left( \mathbf{1}_n^T \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \mathbf{1}_n \right)^{-1} \mathbf{1}_n^T \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \mathcal{L} . \quad (117)$$

Analog zu (74) bei der Parameterschätzung im Kollokationsmodell definieren wir auch hier wieder eine alternative Metrik mit

$$\bar{\Sigma}^{-1} = \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \Sigma^{-1} \quad (118)$$

und können (117) umschreiben zu

$$\mathbf{1}_n \tilde{\mathcal{X}}_{ES} = \mathbf{1}_n \left( \mathbf{1}_n^T \bar{\Sigma}^{-1} \mathbf{1}_n \right)^{-1} \mathbf{1}_n^T \bar{\Sigma}^{-1} \mathcal{L} . \quad (119)$$

Durch Einführung des Projektors

$$\Pi_{S(\mathbf{1}_n)}^{\bar{\Sigma}^{-1}} := \mathbf{1}_n \left( \mathbf{1}_n^T \bar{\Sigma}^{-1} \mathbf{1}_n \right)^{-1} \mathbf{1}_n^T \quad (120)$$

erhält man

$$\mathbf{1}_n \tilde{\mathcal{X}}_{ES} = \left\langle \Pi_{S(\mathbf{1}_n)}^{\bar{\Sigma}^{-1}}, \mathcal{L} \right\rangle_{\bar{\Sigma}^{-1}} . \quad (121)$$

Fasst man diese Erkenntnisse zusammen, so ergibt sich das geschätzte Signal  $\tilde{\mathcal{S}}$  aus (111) übersichtlich zu

$$\tilde{\mathcal{S}} = \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1} \mathcal{L} + \mathbf{1}_n \tilde{\mathcal{X}}_{ES} . \quad (122)$$

beziehungsweise in Skalarproduktschreibweise

$$\tilde{\mathcal{S}} = \left\langle \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}, \mathcal{L} \right\rangle_{\Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}} + \left\langle \Pi_{S(\mathbf{1}_n)}^{\bar{\Sigma}^{-1}}, \mathcal{L} \right\rangle_{\bar{\Sigma}^{-1}} . \quad (123)$$

Zunächst scheint die Metrik  $\Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}$  geometrisch nicht anschaulich zu sein. Durch Einführung der verkürzten Beobachtungen

$$\widetilde{\Delta \mathcal{L}} = \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1} \mathcal{L} = \left\langle \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}, \mathcal{L} \right\rangle_{\Sigma^{-1}} \quad (124)$$

und der Darstellung

$$\tilde{\mathcal{S}} = \left\langle \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}, \widetilde{\Delta \mathcal{L}} \right\rangle_{\Sigma^{-1}} + \left\langle \Pi_{S(\mathbf{1}_n)}^{\bar{\Sigma}^{-1}}, \mathcal{L} \right\rangle_{\bar{\Sigma}^{-1}} \quad (125)$$

ist aber unmittelbar einsichtig, dass die Metrik  $\Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}$  dafür sorgt, dass der Parameteranteil und der konstante Signalanteil aus den Beobachtungen  $\mathcal{L}$  reduziert wird.

Vielfach will man aber auch das Signal ohne den konstanten Anteil des Erwartungswertes des Signal schätzen und erhält somit den variablen Anteil des Signals durch

$$\begin{aligned} \widetilde{\Delta \mathcal{S}} &= \widetilde{\mathcal{S}} - \mathbf{1}_n \widetilde{\mathcal{X}}_{E_S} \\ &= \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1} \mathcal{L} \\ &= \left\langle \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}, \mathcal{L} \right\rangle_{\Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}} \\ &= \left\langle \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}, \widetilde{\Delta \mathcal{L}} \right\rangle_{\Sigma^{-1}} \end{aligned} \quad (126)$$

und somit die Varianz der Schätzung mit

$$\Sigma \{ \widetilde{\Delta \mathcal{S}} \} = \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1} \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} \quad (127)$$

$$= \left\langle \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}, \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} \right\rangle_{\Sigma^{-1} \Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)} \Sigma^{-1}} \cdot \quad (128)$$

Man beachte, dass die Schätzung des Signals unabhängig von der Annahme ob der Erwartungswert des Signals gleich Null ist,  $E_S = 0$  oder ungleich Null ist  $E_S \neq 0$  auf sehr ähnlich Formen führt. Vergleiche (102) und (126) bzw. (104) und (128) für die Varianzen. Lediglich der Projektor ist von  $\Pi_{S^\perp(\mathbf{A})}$  auf  $\Pi_{S^\perp(\mathbf{A}|\mathbf{1}_n)}$  auszutauschen um von einer Form in die andere zu gelangen, wenn man im Falle  $E_S \neq 0$  nur die variablen Anteile des Signals schätzen will.

### 3.3 Schätzung für das Gesamtmodell: Trend und Signal

Zur Schätzung des Gesamtmodells wird der Trend  $\mathbf{A}\widetilde{\mathcal{X}}$  und das Signal  $\widetilde{\mathcal{S}}$  durch

$$\widetilde{\mathcal{L}} = \mathbf{A}\widetilde{\mathcal{X}} + \widetilde{\mathcal{S}} \quad (129)$$

zusammengefasst und ausgeglichene Größen in den Messpunkten  $\widetilde{\mathcal{L}}$  berechnet. Für die weiteren Überlegungen müssen wieder die Fallunterscheidung  $E_S = 0$  bzw.  $E_S \neq 0$  beachten. Zunächst werden wir jedoch das Standardmodell  $E_S = 0$  untersuchen und erst dann die Spezialisierung auf weitere Modelle vornehmen. Wobei wir dann auch auf trendfreie Modelle - Modelle ohne Parameteranteile - übergehen werden, wie sie bei der Prädiktion, Wiener-Kolmogorov-Filterung und dem Kriging angesetzt werden.

**Schätzer für  $\widetilde{\mathcal{L}}$  bei  $E_S = 0$ :** Mit Hilfe von (45) und (101) kann (129) umgeschrieben

$$\widetilde{\mathcal{L}} = \left( \mathbf{A}\widetilde{\mathbf{G}}^T + \widetilde{\mathbf{H}}^T \right) \mathcal{L} \quad (130)$$

und durch Projektoren ausgedrückt werden

$$\widetilde{\mathcal{L}} = \left( \Pi_{S(\mathbf{A})} \Sigma^{-1} + \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(\mathbf{A})} \right) \Sigma^{-1} \mathcal{L}. \quad (131)$$

Bemerkenswert ist, dass sich die Schätzung aus den Anteilen von zwei orthogonalen Vektorräumen zusammensetzt. Während der Vektorraum  $S(\mathbf{A})$  vollständig für die Schätzung genutzt wird, werden für die Signalanteile aus  $S^\perp(\mathbf{A})$  nur Teilbereiche durch Korrelation mit der Kovarianzfunktion ausgewählt. Dadurch wird der Anteil des zufälligen Rauschens  $\mathcal{N}$  unterdrückt.

Die Eigenschaften der Idempotenz und Orthogonalität der Projektoren (Tab. 2) vereinfachen nun die Berechnung der Varianz  $\Sigma \{ \tilde{\mathcal{L}} \}$  durch Varianzfortpflanzung und es ergibt sich

$$\Sigma \{ \tilde{\mathcal{L}} \} = \Pi_{S(A)}^{\Sigma^{-1}} + \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \Sigma^{-1} \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \} . \quad (132)$$

Man erkennt, dass bedingt durch die Orthogonalitäten der Anteile  $A\tilde{\mathcal{X}}$  und  $\tilde{\mathcal{S}}$  die Gesamtvarianz  $\Sigma \{ \tilde{\mathcal{L}} \}$  aus der Summe der Einzelvarianzen  $\Sigma \{ A\tilde{\mathcal{X}} \}$  aus (55) und  $\Sigma \{ \tilde{\mathcal{S}} \}$  aus (103) zusammengesetzt ist. Für die Approximation an beliebigen Punkten (Aufpunkten)  $p \in P$ , mit  $\#P = q$ , kann (129) entsprechend adaptiert werden zu

$$\mathcal{L}_p = A_p \tilde{\mathcal{X}} + \mathcal{S}_p \quad p = 1, \dots, q \quad (133)$$

wobei  $A_p$  den entsprechenden funktionalen Zusammenhang am Aufpunkt  $p$  darstellt. Das geschätzte Signal  $\hat{\mathcal{S}}_p$  errechnet sich analog zu (101) mit

$$\mathcal{S}_p = \Sigma \{ \mathcal{S}_p, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \widetilde{\Delta \mathcal{L}} , \quad p = 1, \dots, q \quad (134)$$

wobei der liegende Vektor  $\Sigma \{ \mathcal{S}_p, \mathcal{S} + \mathcal{N} \}$  die Kovarianzen zwischen dem Aufpunkt  $p$  und allen Messpunkten repräsentiert. Fassen wir nun die Berechnung aller Aufpunkte  $p \in P$  in Vektoren und Matrizen zusammen und kennzeichnen die zusammengefassten Größen mit  $\hat{\mathcal{L}}$ ,  $\hat{A}$  und  $\hat{\mathcal{S}}$  so kann die Berechnung aller approximierten Größen durch

$$\begin{aligned} \hat{\mathcal{L}} &= \hat{A} \tilde{\mathcal{X}} + \hat{\mathcal{S}} \\ &= \hat{A} \tilde{\mathcal{X}} + \Sigma \{ \hat{\mathcal{S}}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \widetilde{\Delta \mathcal{L}} \\ &= \left( \hat{A} (A^T \Sigma^{-1} A)^{-1} A + \Sigma \{ \hat{\mathcal{S}}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \right) \Sigma^{-1} \mathcal{L} \end{aligned} \quad (135)$$

zusammengefasst werden. Die Kovarianzmatrix errechnet sich durch Varianzfortpflanzung mit

$$\Sigma \{ \hat{\mathcal{L}} \} = \hat{A} (A^T \Sigma^{-1} A)^{-1} \hat{A}^T + \Sigma \{ \hat{\mathcal{S}}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(A)}^{\Sigma^{-1}} \Sigma^{-1} \Sigma \{ \mathcal{S} + \mathcal{N}, \hat{\mathcal{S}} \} , \quad (136)$$

wobei wieder die Orthogonalität zwischen der Schätzung von  $\hat{A}\tilde{\mathcal{X}}$  und  $\hat{\mathcal{S}}$  dafür sorgt, dass alle gemischten Anteile entfallen und die Kovarianzen aus der Summe der Einzelkovarianzen gebildet wird.

Nun wollen wir dieses Ergebnis auf alle bisher behandelten Sonderfälle herunterbrechen. In Tab. 3 haben wir alle Ergebnisse zusammengefasst. Dazu wird zunächst eine Ersatzmatrix  $B$  eingeführt und die bisher erhaltenen Formeln für die Schätzung des Gesamtmodells und deren Kovarianzen einfach durch Variation der Ersatzmatrix abgeleitet.

## 4 Modellierung von Kovarianzen

Die Beschreibung der stochastischen Eigenschaften des Signals  $\mathcal{S}$  und des Rauschens  $\mathcal{N}$  erfolgt über die Varianzen/Kovarianzen  $\Sigma \{ \mathcal{S} \}$  bzw.  $\Sigma \{ \mathcal{N} \}$  und werden in den Kovarianzmatrizen  $\Sigma_{\mathcal{S}\mathcal{S}}$  und  $\Sigma_{\mathcal{N}\mathcal{N}}$  zusammengefasst.

Das zufällige Störsignal (Rauschen)  $\mathcal{N}$  wird vielfach als unkorrelierter stochastischer Prozess (“weißes Rauschen”)  $\{ \mathcal{N}(t); t \in \mathbb{R}^d \}$  bzw. diskret für bestimmte Stützstellen  $\{ \mathcal{N}_{t_i}; i \in \mathbb{Z} \}$  angesehen und die Kovarianzmatrix  $\Sigma_{\mathcal{N}\mathcal{N}}$  durch eine mit der Varianz des Rauschens  $\sigma_{\mathcal{N}}^2$  skalierte Einheitsmatrix  $\sigma_{\mathcal{N}}^2 \mathbb{1}_n$  oder eine allgemeine Diagonalmatrix dargestellt, falls das zufällige Störsignal nicht stationär ist. Die

$B := [A] :$	Kollokation, Universal Kriging mit $E_{\mathcal{S}} = 0$
$B := [A   \mathbf{1}_n] :$	Kollokation mit $E_{\mathcal{S}} \neq 0$
$B := [\mathbf{1}_n] :$	Ordinary Kriging
$B := [\emptyset] :$	Prädiktion, Wiener-Kolmogorov-Filterung, Simple Kriging
Filterung:	$\tilde{\mathcal{L}} = \left( \Pi_{S(B)}^{\Sigma^{-1}} + \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(B)}^{\Sigma^{-1}} \right) \Sigma^{-1} \mathcal{L}$ $\Sigma \{ \tilde{\mathcal{L}} \} = \Pi_{S(B)}^{\Sigma^{-1}} + \Sigma \{ \mathcal{S}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(B)}^{\Sigma^{-1}} \Sigma^{-1} \Sigma \{ \mathcal{S} + \mathcal{N}, \mathcal{S} \}$
Prädiktion:	$\hat{\mathcal{L}} = \left( \hat{B} \left( B^T \Sigma^{-1} B \right)^{-1} B + \Sigma \{ \hat{\mathcal{S}}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(B)}^{\Sigma^{-1}} \right) \Sigma^{-1} \mathcal{L}$ $\Sigma \{ \hat{\mathcal{L}} \} = \hat{B} \left( B^T \Sigma^{-1} B \right)^{-1} \hat{B}^T + \Sigma \{ \hat{\mathcal{S}}, \mathcal{S} + \mathcal{N} \} \Sigma^{-1} \Pi_{S^\perp(B)}^{\Sigma^{-1}} \Sigma^{-1} \Sigma \{ \mathcal{S} + \mathcal{N}, \hat{\mathcal{S}} \}$

**Table 3:** Schätzung des Gesamtmodells in den unterschiedlichsten Variationen. Durch Variation der Matrix  $B$  können mit den gleichen Formeln sowohl die Filterung als auch die Prädiktion an Aufpunkten bewerkstelligt werden. Die Matrix  $\hat{B}$  für die Prädiktion ist analog zu  $B$  aufzubauen und abhängig vom Modell. Wenn die Matrix  $B$  keine Spalten hat und damit der Spaltenraum von  $B$  die leere Menge  $\emptyset$  ist, so entartet die Projektionsmatrix zur Nullmatrix  $\Pi_{S(\emptyset)}^{\Sigma^{-1}} = \mathbf{O}_n$  und die orthogonale Projektionsmatrix zur Kovarianzmatrix  $\Pi_{S^\perp(\emptyset)}^{\Sigma^{-1}} = \Sigma$ .

Signale  $\mathcal{S}$  sind hingegen korreliert und die Einträge in der Kovarianzmatrix  $\Sigma_{\mathcal{S}\mathcal{S}}$  werden über (Auto)Kovarianzfunktionen erzeugt. Für einen stationären Prozess  $\{ \mathcal{S}(\mathbf{t}); \mathbf{t} \in \mathbb{R}^d \}$  bzw.  $\{ \mathcal{S}_{t_i}; i \in \mathbb{Z} \}$  sind die Autokovarianzen  $\Sigma \{ \mathcal{S}(\mathbf{t}_i), \mathcal{S}(\mathbf{t}_j) \}$  bzw.  $\Sigma \{ \mathcal{S}_{t_i}, \mathcal{S}_{t_j} \}$  wegen der Forderung (9) unabhängig vom Ort und können zum Beispiel durch Autokovarianzfunktionen (Kap. 4.1) oder diskrete kovarianzstationäre Prozesse beschrieben werden. Zunächst werden wir in Kap. 4.1 die mathematischen Grundlagen zu Kovarianzfunktionen betrachten und dann spezielle finite Kovarianzfunktionen diskutieren, die bedingt durch den endlichen Träger auf dünnbesetzte Kovarianzmatrizen führen, die auch bei einer großen Anzahl von Messungen noch effizient berechenbar sind. Im folgenden Kapitel 4.3 diskutieren wir einen Zugang zur Modellierung von korrelierten Messreihen über diskrete kovarianzstationäre Prozesse. Der Übergang und Zusammenhang zu Kovarianzfunktionen wird über das *Magische Quadrat* hergestellt.

## 4.1 Kovarianzfunktion

Die Autokovarianzfunktion  $\gamma(t_{ij})$  ist analog zu (9) definiert durch

$$\Sigma \{ \mathcal{S}(\mathbf{t}_i), \mathcal{S}(\mathbf{t}_j) \} = \Sigma \{ \mathcal{S}(\mathbf{t}_i - \mathbf{t}_j), \mathcal{S}(\mathbf{t}_j - \mathbf{t}_j) \} := \gamma(\mathbf{t}_j - \mathbf{t}_i), \quad (137)$$

wobei  $\mathbf{t}_i$  und  $\mathbf{t}_j$  Punktkoordinaten in einem Gebiet  $\mathcal{A}$  von  $\mathbb{R}^d$  darstellen und somit  $\mathbf{t}_j - \mathbf{t}_i$  die Koordinatendifferenzen („lags“) kennzeichnen. Vereinfacht gilt für skalare Größen wie Zeitdifferenzen  $t_{ij} = t_j - t_i$ , wobei  $t_i, t_j \in \mathbb{R}$  konkrete Zeitpunkte  $i, j = 1, \dots, n$  kennzeichnen. Bei mehrdimensionalen Anwendungen kommen vielfach auch isotrope Kovarianzfunktionen zum Einsatz, wo die Koordinatendifferenzen  $\mathbf{t}_j - \mathbf{t}_i$  durch eine geeignete Norm  $\| \mathbf{t}_j - \mathbf{t}_i \| = t_{ij}$  ersetzt werden und somit die Kovarianzfunktion nur eine Funktion des Distanz  $t_{ij}$  ist. In den folgenden Betrachtungen wollen wir aber zunächst noch

den allgemeinen nicht isotropen Fall verfolgen und kommen dann am Ende dieses Abschnittes auf die Verwendung von isotropen Funktionen zurück.

Nach (Brockwell und Davis, 1991, S. 26, Proposition 1.5.1) weisen Autokovarianzfunktionen  $\gamma(\mathbf{t}_j - \mathbf{t}_i)$  eines stationären stochastischen Prozesses folgende Eigenschaften auf:

$$\bullet \gamma(\mathbf{0}_d) \geq 0, \quad (138)$$

$$\bullet |\gamma(\mathbf{t}_j - \mathbf{t}_i)| \leq \gamma(\mathbf{0}_d) \quad \text{für alle } \mathbf{t}_i, \mathbf{t}_j \in \mathbb{R}^d, \quad (139)$$

$$\bullet \gamma(\mathbf{t}_j - \mathbf{t}_i) = \gamma(-(\mathbf{t}_j - \mathbf{t}_i)) \quad \text{und} \quad (140)$$

$$\bullet \gamma(\mathbf{t}_j - \mathbf{t}_i) \text{ ist eine positiv semidefinite Funktion.} \quad (141)$$

Eine Funktion  $\gamma(\mathbf{t}_j - \mathbf{t}_i)$  ist positiv semidefinit, wenn

$$\sum_{i,j=1}^n \gamma(\mathbf{t}_i - \mathbf{t}_j) z_i z_j \geq 0 \quad (142)$$

für jede Wahl von  $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbb{R}^d$ ,  $z_1, \dots, z_n \in \mathbb{R}$  gültig ist, wobei das Gleichheitszeichen für **se**-**m**idefinite Funktionen steht. Aus dieser Definition folgt unmittelbar, dass eine Linearkombination von positiv semidefiniten Funktionen  $\gamma_k(\mathbf{t})$  mit Koeffizienten  $c_k \geq 0$

$$\gamma(\mathbf{t}) = \sum_k c_k \gamma_k(\mathbf{t}) \quad (143)$$

wieder auf eine positiv semidefinite Funktion führt. Somit sind also die Addition bzw. die Multiplikation mit einem positiven Skalar erlaubte Operationen um aus positiv semidefiniten Funktionen weitere Funktionen abzuleiten, die ebenfalls positiv semidefinit sind. Zur Ableitung weiterer Rechenregeln spielt die Fouriertransformierte einer integrierbaren, kontinuierlichen, skalaren Funktion  $f(\mathbf{t})$  eine besondere Rolle. Diese ist definiert durch

$$f(\mathbf{t}) \xleftrightarrow{\mathcal{F}} F(\boldsymbol{\nu}) \quad \text{mit} \quad \begin{cases} F(\boldsymbol{\nu}) := \mathcal{F}\{f(\mathbf{t})\} & := \int_{-\infty}^{\infty} f(\mathbf{t}) e^{-i2\pi\langle\boldsymbol{\nu},\mathbf{t}\rangle} d\mathbf{t} \\ f(\mathbf{t}) := \mathcal{F}^{-1}\{F(\boldsymbol{\nu})\} & := \int_{-\infty}^{\infty} F(\boldsymbol{\nu}) e^{i2\pi\langle\boldsymbol{\nu},\mathbf{t}\rangle} d\boldsymbol{\nu} \end{cases} \quad (144)$$

wobei  $\langle \cdot, \cdot \rangle$  das Skalarprodukt benennt,  $\mathbf{t}, \boldsymbol{\nu} \in \mathbb{R}^d$  repräsentieren kontinuierliche Vektoren im Orts- bzw. Frequenzbereich und  $F(\boldsymbol{\nu})$  kennzeichnet die Fouriertransformierte von  $f(\mathbf{t})$ , die durch die Fouriertransformation in  $\mathbb{R}^d$  gebildet wird. Das Theorem von Bochner (Bochner, 2005, S. 55) besagt nun, dass für eine positiv definite Funktion  $f(\mathbf{t})$  gilt, dass die Fourier-Transformierte  $F(\boldsymbol{\nu})$  nur nicht negative Werte aufweisen muss.

Dieses Theorem kann über die inverse Fourier Transformation  $\mathcal{F}^{-1}\{\cdot\}$  von (142) für  $\mathbf{t} \in \mathbb{R}^d$  bewiesen werden. Unter Beachtung des Verschiebesatzes für Fourier-Integrale

$$f(\mathbf{t} - \mathbf{t}_0) = \mathcal{F}^{-1}\left\{F(\boldsymbol{\nu})e^{-i2\pi\langle\boldsymbol{\nu},\mathbf{t}_0\rangle}\right\} = \int_{-\infty}^{\infty} F(\boldsymbol{\nu})e^{i2\pi\langle\boldsymbol{\nu},\mathbf{t}-\mathbf{t}_0\rangle} d\boldsymbol{\nu} \quad (145)$$



kann (142) umgeformt werden zu

$$\begin{aligned}
\sum_{j,k=1}^n f(\mathbf{t}_j - \mathbf{t}_k) z_j z_k &= \sum_{j,k=1}^n \left( \int_{-\infty}^{\infty} e^{i2\pi \langle \boldsymbol{\nu}, \mathbf{t}_j - \mathbf{t}_k \rangle} F(\boldsymbol{\nu}) d\boldsymbol{\nu} \right) z_j z_k \\
&= \int_{-\infty}^{\infty} \left( \sum_{j=1}^n e^{i2\pi \langle \boldsymbol{\nu}, \mathbf{t}_j \rangle} z_j \right) \left( \sum_{k=1}^n e^{-i2\pi \langle \boldsymbol{\nu}, \mathbf{t}_k \rangle} z_k \right) F(\boldsymbol{\nu}) d\boldsymbol{\nu} \\
&= \int_{-\infty}^{\infty} \left| \sum_{j=1}^n e^{i2\pi \langle \boldsymbol{\nu}, \mathbf{t}_j \rangle} z_j \right|^2 F(\boldsymbol{\nu}) d\boldsymbol{\nu} \geq 0, \tag{146}
\end{aligned}$$

wobei durch die quadratische Form der Summe gezeigt ist, dass für eine positiv definite Funktion  $f(\mathbf{t})$  alle  $F(\boldsymbol{\nu}) d\boldsymbol{\nu} \geq 0$  sein müssen (Stewart, 1976). Diese Eigenschaft kann man nutzen um auch die Multiplikation von zwei positiv semidefiniten Funktionen  $g(\mathbf{t}) \cdot h(\mathbf{t})$  bzw. die Faltung  $g(\mathbf{t}) * h(\mathbf{t})$  als weitere mögliche Operationen zu identifizieren, die es erlauben positiv semidefinite Funktionen zu erzeugen. Dies folgt unmittelbar aus dem Faltungstheorem (Kreyszig, 1993, Kap. 10.11)

$$g(\mathbf{t}) h(\mathbf{t}) \xleftrightarrow{\mathcal{F}} G(\boldsymbol{\nu}) * H(\boldsymbol{\nu}) \quad \text{bzw.} \quad g(\mathbf{t}) * h(\mathbf{t}) \xleftrightarrow{\mathcal{F}} G(\boldsymbol{\nu}) H(\boldsymbol{\nu}) \tag{147}$$

mit  $G(\boldsymbol{\nu}) = \mathcal{F}\{g(\mathbf{t})\} \geq 0$  und  $H(\boldsymbol{\nu}) = \mathcal{F}\{h(\mathbf{t})\} \geq 0$ . Zuletzt sei noch festgehalten, dass die Faltung einer beliebigen Funktion mit sich selbst

$$\int_{-\infty}^{\infty} h(\mathbf{z}) h(\mathbf{t} - \mathbf{z}) d\mathbf{z} = h(\mathbf{t}) * h(\mathbf{t}) \tag{148}$$

wegen

$$h(\mathbf{t}) * h(\mathbf{t}) \xleftrightarrow{\mathcal{F}} H(\boldsymbol{\nu}) H(\boldsymbol{\nu}) \geq 0 \tag{149}$$

die Konstruktion von positiv semidefiniten Funktionen erlaubt. Zusammenfassend können somit folgende Rechenvorschriften festgehalten werden, um positiv semidefinite Funktionen zu erzeugen:

- die Addition/Multiplikation von zwei positiv semidefiniten Funktionen,
- die Multiplikation einer semidefiniten Funktion mit einem positiven Skalar und
- die Faltung einer Funktion mit sich selbst.

Ausgehend von einer positiv semidefiniten Kovarianzfunktion  $\gamma(\mathbf{t}_j - \mathbf{t}_i)$  kann nun (142) verwendet werden um eine positiv semidefinite Kovarianzmatrix  $\boldsymbol{\Sigma}$  aufzubauen. Aus (142) folgt unmittelbar, dass eine quadratische  $n \times n$  Matrix  $\boldsymbol{\Sigma}$  gebildet durch die Vorschrift  $\sigma_{ij} = \gamma(\mathbf{t}_j - \mathbf{t}_i)$  aus der positive semidefiniten Funktion  $\gamma(\cdot)$  wegen

$$\sum_{i,j=1}^n \gamma(\mathbf{t}_j - \mathbf{t}_i) z_i z_j = \sum_{i,j=1}^n \sigma_{ij} z_i z_j = \mathbf{z}^T \boldsymbol{\Sigma} \mathbf{z} \geq 0 \quad \text{mit} \quad \mathbf{z} = [z_1, \dots, z_n]^T \tag{150}$$

eine positiv semidefinite Matrix repräsentiert. Diese Matrix ist wegen (140) bzw.

$$\sigma_{ij} = \gamma(\mathbf{t}_i - \mathbf{t}_j) = \gamma(\mathbf{t}_j - \mathbf{t}_i) = \sigma_{ji} \tag{151}$$

symmetrisch. Zusammenfassend kann also festgehalten werden, dass aufgrund des Theorems von Bochner (Bochner, 2005, S. 55) die symmetrische Matrix  $\boldsymbol{\Sigma}$  mit den Eintragungen  $\sigma_{ij} = \gamma(\mathbf{t}_i - \mathbf{t}_j) = \gamma(\mathbf{t}_j - \mathbf{t}_i)$ ;

$t_i, t_j \in \mathcal{A}$ , positiv semidefinit für alle Punktanordnungen in  $\mathcal{A}$  von  $\mathbb{R}^d$  ist, falls die Fouriertransformierte  $F(\nu)$  der Funktion  $f(t) : \mathbb{R}^d \rightarrow \mathbb{R}$  nicht negativ ist (Gaspari und Cohn, 1999, Theorem 2.10).

Das Bochner Theorem garantiert somit auch die positive Semidefinitheit für jeden Unterraum  $\mathcal{A} \subset \mathbb{R}^d$ . Daher führt eine positive semidefinite Funktion in  $\mathbb{R}^d$  auf positiv semidefinite Matrizen für jeden Unterraum von  $\mathbb{R}^d$ , z.B. auf einer Linie, in einer Ebenen, auf einer Kugel (Gaspari und Cohn, 1999).

Für eindimensionale Prozesse kann somit auf viele positiv definite Funktionen aus höherdimensionalen Räumen zugegriffen werden. Positiv definite Funktionen in  $\mathbb{R}^n$  (e.g.  $e^{-a|r|}$ ,  $e^{-ar^2}$ ) können ebenso eingesetzt werden wie Bessel-Funktionen erster Art  $J_1(ar)/ar$  or  $J_0(ar)$  aus  $\mathbb{R}^2$  oder  $\cos(\omega t)$  und  $\text{sinc}(\omega t)$  aus  $\mathbb{R}^1$  (Mussio, 1984). In diesem Zusammenhang sei auch speziell darauf hingewiesen, dass auch für isotrope Funktionen in  $\mathbb{R}^d$  die positive Semidefinitheit der Kovarianzfunktionen im entsprechenden Koordinatenraum  $\mathbb{R}^d$  gewährleistet sein muss. Die irrige Annahme, dass die durch die Norm abgeleiteten Distanzen nur skalare Größen darstellen und damit eine positive semidefinite Funktion in  $\mathbb{R}^1$  hinreichend ist, kann durch Gegenbeispiele widerlegt werden (Sansò, 2015 - persönliche Kommunikation).

In der physikalischen Geodäsie sind Kovarianzfunktionen für harmonische Funktionen im oberen Halbraum und für globale Anwendungen im Außenraum der Kugel von großer Bedeutung. Basierend auf den grundlegenden Arbeiten von Heiskanen und Moritz (1967); Hirvonen (1962); Meissl (1971); Moritz (1976); Tscherning (1972); Tscherning und Rapp (1974) findet man bei Moritz (1980) eine umfassende Darstellung. Softwarepakete wie zum Beispiel GRAVSOFTE (Tscherning et al., 1994) unterstützen die Anwendung von Kollokationsverfahren in der täglichen Praxis der Erschwerfeldbestimmung.

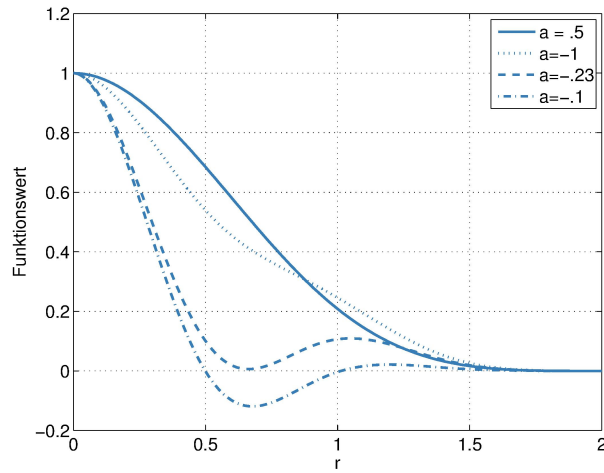
Automatisierte Messungen führen sehr rasch auf große Datenmengen und somit auf große Kovarianzmatrizen. Um mit diesen großen Datenmengen numerisch zurecht zu kommen, bieten sich mehrere Zugänge speziell an. Bei äquidistanter Datenverteilung kann die Toeplitz-Struktur der Kovarianzmatrix genutzt werden um speziell adaptierte Verfahren zu entwickeln, die auch mit sehr großen Datenmengen noch verwendet werden können (Eren (1982), Bottoni und Barzaghi (1993), Schuh (1996), Sansò und Tscherning (2003)). Im Abschnitt 4.3 werden wir Kovarianzmatrizen durch lineare diskrete zeitinvariante Filter modellieren und damit eine weitere Möglichkeiten zur effizienten Verarbeitung von sehr langen Zeitreihen aufzeigen. Liegt eine unregelmäßige Abtastung vor, so können mit *finiten Kovarianzfunktionen* (= Kovarianzfunktionen mit endlichem Träger) dünnbesetzte Kovarianzmatrizen erzeugt werden, die auch für sehr große Systeme noch effizient lösbar sind. Im folgenden Abschnitt 4.2 wird das Grundprinzip zur Konstruktion von finiten Kovarianzfunktionen erläutert, Beispiele für finite Kovarianzfunktionen in  $\mathbb{R}^1$  und  $\mathbb{R}^3$  aufgezeigt und anschließend finite Kovarianzfunktionen in  $\mathbb{R}^2$  entwickelt.

## 4.2 Finite Kovarianzfunktionen

Als *finite Kovarianzfunktionen* bezeichnet man Kovarianzfunktionen, die nur in einem bestimmten Bereich Funktionswerte ungleich Null aufweisen und außerhalb dieses Bereiches exakt Null sind. Diese Funktionen haben also einen endlichen Träger und führen daher auf dünnbesetzte Kovarianzmatrizen.

Die in der Praxis oft verwendete Methode (sehr) kleine Funktionswerte einer abklingenden Kovarianzfunktion durch Nullen zu ersetzen, entspricht der Multiplikation der originalen Kovarianzfunktion mit einer Rechtecksfunktion (*Box-Car-Funktion*). Die Fouriertransformierte einer Rechtecksfunktion bildet die sinc-Funktion, die auch negative Werte aufweist. Somit kann die positive Definitheit der so modifizierten Kovarianzfunktion nicht mehr gewährleistet werden. Die resultierende Kovarianzmatrix ist zwar dünn besetzt, aber nicht notwendigerweise auch positiv definit.

Für den eindimensionalen Fall bilden die Basissplines  $S1$  (*Hütchen-Funktion*, Spline erster Ordnung) und  $S3$  (kubischer Spline) mögliche positiv semidefinite Funktionen (Sansò und Schuh, 1987). Die positive Semidefinitheit ist durch die Definitionen von  $S1 := S0 * S0$  und  $S3 := S1 * S1$  nach (149)



**Abb. 3:** Gaspari-Cohn Funktion:  $f_{GC}(r, a, R)$  mit endlichem Träger 2 ( $R=1$ ). Variationen des Formparameters  $a$ .

gewährleistet, da beide Basissplines durch Faltung der Funktionen  $S_0$  (Rechteck-Funktion, Spline nullter Ordnung) bzw.  $S_1$  mit sich selbst definiert werden können. Ähnliche Überlegungen gelten somit für alle Basissplines mit ungeraden Ordnungen.

Gaspari und Cohn (1999) berechnen eine stückweise definierte Funktion fünfter Ordnung  $f_{GC}(r, a, R)$  in den Parametern  $a$  und  $R$  durch eine Faltung in  $\mathbb{R}^3$  der stückweisen linearen, stetigen Funktion

$$h(r, a, R) = \begin{cases} 2(a-1)r/R + 1 & : 0 \leq r \leq R/2 \\ 2a(1-r/R) & : R/2 \leq r \leq R \\ 0 & : R \leq r \end{cases} \quad (152)$$

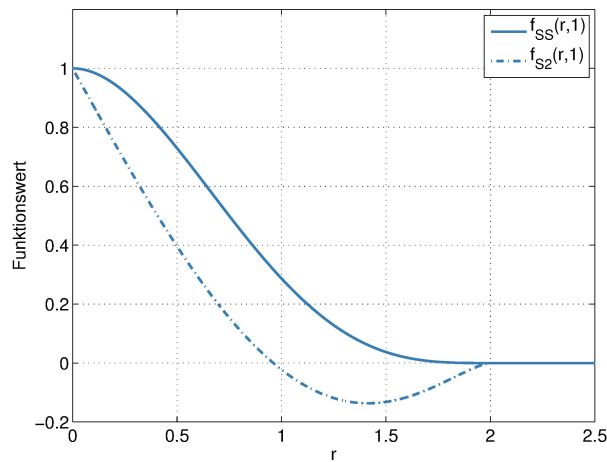
mit sich selbst. Das Faltungstheorem (Papoulis, 1977) garantiert auch für höhere Dimensionen die positive Semidefinitheit der resultierenden Funktion  $f_{GC}(r, a, R)$ , wobei  $r$  die Euklidische Distanz  $\|\mathbf{t}_i - \mathbf{t}_j\|$  zwischen den Punktkoordinaten  $\mathbf{t}_i, \mathbf{t}_j \in \mathbb{R}^3$  bezeichnet.  $R$  definiert die halbe Länge des endlichen Trägers. Der Formparameter  $a \in \mathbb{R}$ , der die Steigungen der linearen Ausgangsfunktionen festlegt, erlaubt eine flexible Steuerung der Gestalt der resultierenden positiv semidefiniten Funktion. Abb. 3 zeigt den Verlauf für unterschiedliche Formparameter auf. Die umfangreichen Formeln des stückweise definierten Polynoms fünfter Ordnung  $f_{GC}(z, a, R)$  sind bei Gaspari et al. (2006) [(33), (C.1) und (C.2)] zu finden und werden auch bei Moreaux (2008)[(41)-(43)] wiedergegeben. Durch die Wahl der Parameter  $a$  und  $R$  kann die Gestalt der positiv semidefiniten Funktion sehr variabel verändert werden, sodass eine große Freiheit zur Anpassung an die empirische Kovarianzen ermöglicht wird.

Sansò und Schuh (1987) verwenden die parabolische Funktion

$$h(r, R) = \begin{cases} R^2 - r^2 & : 0 \leq r \leq R \\ 0 & : R \leq r \end{cases} \quad (153)$$

als Ausgangsfunktion, wobei  $R$  den Träger festlegt. Durch Faltung in  $\mathbb{R}^2$  der Ausgangsfunktion mit sich selbst ergibt sich

$$f_{SS}(r, R) = \begin{cases} \frac{1}{3}R^6\pi - \frac{1}{2}R^4r^2\pi + \\ + \frac{1}{3}(R^4r + \frac{4}{3}R^2r^3 - \frac{1}{12}r^5) \sqrt{R^2 - (\frac{r}{2})^2} + \\ + (R^4r^2 - \frac{2}{3}R^6) \arcsin \frac{r}{2R} & : 0 \leq r < 2R \\ 0 & : 2R \leq r \end{cases} \quad (154)$$



**Abb. 4:** Sansò-Schuh Funktionen:  $\bar{f}_{SS}(r, R)$  aus (155) und  $\bar{f}_{S2}(r, R)$  aus (166) mit endlichem Träger 2 ( $R=1$ )

und bildet somit eine positiv semidefinite Funktion in  $\mathbb{R}^2$ , wobei der endliche Träger durch den zweifachen Wert des Parameters  $R$  festgelegt wird. Die skalierte Version

$$\bar{f}_{SS} = f_{SS}(r, R)/f_{SS}(0, R) \quad (155)$$

ist in Figur 4 dargestellt. Diese Funktion  $\bar{f}_{SS}(r, R)$  hat eine geringere Variabilität als  $f_{GC}$ , aber eine relativ einfache mathematische Darstellung. Durch multiplikative Verknüpfung mit anderen positiv definiten Funktionen (siehe Kap. 4.1) kann aber wieder eine große Variabilität erreicht werden. Die multiplikative Kombination mit den Exponentialfunktionen, e.g.  $e^{-a|r|}$  oder  $e^{-ar^2}$ , ermöglicht beispielsweise eine bessere Approximation der Krümmung am Ursprung. Oszillationen können durch Kombination mit den Bessel-Funktionen  $J_1(ar)/ar$  or  $J_0(ar)$  erreicht werden (Schuh, 1989).

Allerdings können mit der Funktion  $\bar{f}_{SS}(r, R)$  ohne Hinzunahme der Bessel-Funktionen zunächst keine negative Korrelationen modelliert werden. Hier ist die Gaspari-Cohn Basisfunktion viel flexibler. Zum Beispiel erlaubt die Wahl eines negativen Formparameters  $a$  (z.B.  $a = -0.1$  in Fig. 3) die Modellierung von negativen Korrelationen, die in der Praxis oft vorkommen.

Um diese Schwäche der Funktion  $\bar{f}_{SS}(r, R)$  aus (154) zu beheben, kann man das Differentiationstheorem der Fouriertransformation nutzen (Kreyszig, 1993, S. 616). In  $\mathbb{R}^1$  ergibt sich unmittelbar, dass die Fourier-Transformierte der zweiten Ableitung einer Funktion  $\mathcal{F}\{g''(t)\}$  direkt aus der Fourier-Transformierten der Funktion  $G(\nu) = \mathcal{F}\{g(t)\}$  abgeleitet werden kann, denn es gilt

$$g''(t) \xleftrightarrow{\mathcal{F}} -\nu^2 G(\nu) . \quad (156)$$

Wenn  $g(t)$  ein positives Spektrum aufweist und somit eine positiv definite Funktion darstellt, dann ist  $-g''(t)$  auch positiv definit. Diese Eigenschaft kann genutzt werden um positiv semidefinite Funktionen zu erzeugen, indem man die negative zweite Ableitung benutzt.

In Koch et al. (2010) wird dieser Ansatzes für die Konstruktion einer isotropen, finiten Kovarianzfunktion in  $\mathbb{R}^2$  benutzt. Zunächst erfolgt der Übergang von der zweidimensionalen Fourier-Transformation

$$g(x, y) \xleftrightarrow{\mathcal{F}} G(\nu, \mu) := \mathcal{F}\{g(x, y)\} := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) e^{-i2\pi(x\nu+y\mu)} dx dy , \quad (157)$$

auf die Hankel-Transformation

$$g(r) \xleftrightarrow{\mathcal{H}} \bar{G}(\rho) := \int_0^{\infty} r g(r) J_0(\rho r) dr, \quad (158)$$

indem die Funktion  $g(x, y)$  durch die isotrope Funktion  $g(r)$  mit  $r = \sqrt{x^2 + y^2}$  substituiert wird.  $J_0(x)$  kennzeichnet die Bessel-Funktion erster Art nullter Ordnung und ist definiert durch

$$J_0(x) := \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i2\pi x \cos \phi} d\phi. \quad (159)$$

Der Zusammenhang zwischen der zwei-dimensionalen Fourier-Transformierten  $G(\nu, \mu)$  und der Hankel-Transformierten  $\bar{G}(\rho)$  mit  $\rho = \sqrt{\nu^2 + \mu^2}$  kann durch

$$G(\nu, \mu) = 2\pi \bar{G}(\sqrt{\nu^2 + \mu^2}) = \bar{G}(\rho) \quad (160)$$

(Papoulis, 1977, S. 280, Theorem 1) hergestellt werden, wodurch sofort ersichtlich ist, dass für positiv definite Funktionen in  $\mathbb{R}^2$  aus  $G(\nu, \mu) \geq 0$  unmittelbar  $\bar{G}(\rho) \geq 0$  folgt. Der Zusammenhang zwischen der ersten und zweiten Ableitung der radial symmetrischen Funktion  $g(r)$  in  $\mathbb{R}^2$  und der korrespondierenden Hankel-Transformierten  $\bar{G}(\rho)$  ist nach Papoulis (1977, p. 283, eq. (8-78)) durch

$$\frac{d^2 g(r)}{dr^2} + \frac{1}{r} \frac{dg(r)}{dr} \xleftrightarrow{\mathcal{H}} -\omega^2 \bar{G}(\rho) \quad \text{bzw.} \quad g''(r) + \frac{1}{r} g'(r) \xleftrightarrow{\mathcal{H}} -\omega^2 \bar{G}(\rho) \quad (161)$$

gegeben. Aus  $\bar{G}(\rho) \geq 0$  für positive semidefinite Funktionen  $g(r)$  kann unmittelbar gefolgert werden, dass der Ausdruck links vom Gleichheitszeichen eine negativ definite Funktion kennzeichnet. Dies kann nun genutzt werden um aus der positiv semidefiniten Funktion  $f_{SS}(r, R)$  (154) durch Bildung der ersten und zweiten Ableitung eine positiv semidefinite Funktion in  $\mathbb{R}^2$  zu konstruieren.

Die Differentiation von  $f_{SS}(r, R)$  (154) nach  $r$  führt auf

$$f'_{SS}(r, R) = \frac{\partial g(r, R)}{\partial r} = \begin{cases} -R^4 r \pi + 2R^4 r \arcsin \frac{r}{2R} + \\ \quad + \left( \frac{5}{3} R^2 r^2 - \frac{1}{6} r^4 \right) \sqrt{R^2 - \left( \frac{r}{2} \right)^2} & : \quad 0 \leq r \leq 2R \\ 0 & : \quad 2R \leq r. \end{cases} \quad (162)$$

Die zweite Ableitung ergibt

$$f''_{SS}(r, R) = \frac{\partial^2 g(r, R)}{\partial r^2} = \begin{cases} -R^4 \pi + 2R^4 \arcsin \frac{r}{2R} + \\ \quad + \left( \frac{13}{3} R^2 r - \frac{5}{6} r^3 \right) \sqrt{R^2 - \left( \frac{r}{2} \right)^2} & : \quad 0 \leq r \leq 2R \\ 0 & : \quad 2R \leq r. \end{cases} \quad (163)$$

Mit (163) und (162) kann nun die linke Seite von (161) ausgewertet werden

$$\frac{\partial^2 g(r, R)}{\partial r^2} + \frac{1}{r} \frac{\partial g(r, R)}{\partial r} = -2R^4 \pi + 4R^4 \arcsin \frac{r}{2R} + (6R^2 r - r^3) \sqrt{R^2 - \left( \frac{r}{2} \right)^2} \quad (164)$$

und die positive semidefinite Funktion  $f_{S2}(r, R)$  aus der negativen Form der rechten Seite von (164)

$$f_{S2}(r, R) := 2R^4 \pi - 4R^4 \arcsin \frac{r}{2R} - (6R^2 r - r^3) \sqrt{R^2 - \left( \frac{r}{2} \right)^2} \quad (165)$$

abgeleitet werden. Die skalierte Funktion

$$\bar{f}_{S2} = f_{S2}(r, R)/f_{S2}(0, R) \quad (166)$$

ist in Fig. 4 dargestellt. Mit  $\bar{f}_{SS}(r, R)$  aus (155) und  $\bar{f}_{S2}(r, R)$  aus (166) stehen nun zwei positiv semidefinite, finite Funktionen in  $\mathbb{R}^2$  zur Verfügung, die eine flexible Modellierung unterschiedlichster Kovarianzfunktionen erlauben. Erste Anwendungen dieser Funktionen sind bei Koch et al. (2010) zu finden.

Die zuvor dargestellten Ableitungen  $f'_{SS}(r, R)$  (162) und  $f''_{SS}(r, R)$  (163) können auch zur Berechnung der Kovarianzen des Gradienten

$$\nabla \mathcal{S}(\mathbf{t}) = \begin{bmatrix} \frac{\partial \mathcal{S}(\mathbf{t})}{\partial x} \\ \frac{\partial \mathcal{S}(\mathbf{t})}{\partial y} \end{bmatrix} := \begin{bmatrix} \mathcal{S}_x(\mathbf{t}) \\ \mathcal{S}_y(\mathbf{t}) \end{bmatrix} \quad (167)$$

des kontinuierlichen Prozesses  $\{\mathcal{S}(\mathbf{t}); \mathbf{t} \in \mathbb{R}^2\}$  verwendet werden. Durch Anwendungen des Kovarianzfortsetzungssatzes für lineare Funktionale (Moritz, 1980, (11-14)) können aus der isotropen Kovarianzfunktion  $f_{SS}(r, R)$  auch die Kovarianzfunktionen für die Gradienten hergeleitet werden. Nach (Moritz, 1973, S. 66) und (Moritz, 1980, S. 171-173) ergeben sich die Autokovarianzfunktion für die Funktion bzw. die partiellen Ableitungen  $\mathcal{S}_x(\mathbf{t})$  und  $\mathcal{S}_y(\mathbf{t})$  mit

$$\Sigma \{\mathcal{S}(\mathbf{t}_i), \mathcal{S}(\mathbf{t}_j)\} = f_{SS}(r_{ij}, R) \quad (168)$$

$$\Sigma \{\mathcal{S}_x(\mathbf{t}_i), \mathcal{S}(\mathbf{t}_j)\} = -f'_{SS}(r_{ij}, R) \cos \alpha_{ij} = -\Sigma \{\mathcal{S}(\mathbf{t}_i), \mathcal{S}_x(\mathbf{t}_j)\} \quad (169)$$

$$\Sigma \{\mathcal{S}_y(\mathbf{t}_i), \mathcal{S}(\mathbf{t}_j)\} = -f'_{SS}(r_{ij}, R) \sin \alpha_{ij} = -\Sigma \{\mathcal{S}(\mathbf{t}_i), \mathcal{S}_y(\mathbf{t}_j)\} \quad (170)$$

$$\Sigma \{\mathcal{S}_x(\mathbf{t}_i), \mathcal{S}_x(\mathbf{t}_j)\} = -f''_{SS}(r_{ij}, R) \cos^2 \alpha_{ij} - \frac{1}{r_{ij}} f'_{SS}(r_{ij}, R) \sin^2 \alpha_{ij} \quad (171)$$

$$\Sigma \{\mathcal{S}_x(\mathbf{t}_i), \mathcal{S}_y(\mathbf{t}_j)\} = -[f''_{SS}(r_{ij}, R) - \frac{1}{r_{ij}} f'_{SS}(r_{ij}, R)] \sin \alpha_{ij} \cos \alpha_{ij} = \Sigma \{\mathcal{S}_y(\mathbf{t}_i), \mathcal{S}_x(\mathbf{t}_j)\} \quad (172)$$

$$\Sigma \{\mathcal{S}_y(\mathbf{t}_i), \mathcal{S}_y(\mathbf{t}_j)\} = -f''_{SS}(r, R) \sin^2 \alpha - \frac{1}{r_{ij}} f'_{SS}(r_{ij}, R) \cos^2 \alpha_{ij} \quad (173)$$

wobei  $\alpha$  den Richtungswinkel

$$\cos \alpha_{ij} = \frac{x_j - x_i}{r_{ij}} \quad \text{und} \quad \sin \alpha_{ij} = \frac{y_j - y_i}{r_{ij}} \quad (174)$$

und  $r$  die euklidische Entfernung zwischen den Punkten  $\mathbf{t}_i$  und  $\mathbf{t}_j$

$$r_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (175)$$

kennzeichnet. Somit können sowohl direkte Messungen der Signale  $\mathcal{S}(\mathbf{t})$  als auch Messungen von Ableitungen  $\mathcal{S}_x(\mathbf{t})$  bzw.  $\mathcal{S}_y(\mathbf{t})$  in einem Schätzprozess konsistent verarbeitet werden. Die resultierende Kovarianzmatrix ist bei Einsatz der finiten Kovarianzfunktionen dünn besetzt.

### 4.3 Diskreter kovarianzstationärer Prozess

Das stochastische Verhalten einer stationären Folge von Zufallsvariablen  $\{\mathcal{S}_t\}_{\Delta t}$  kann über lineare zeitinvariante stochastische Prozesse modelliert werden. Das zufällige Signal  $\mathcal{S}_t$  zum Zeitpunkt  $t$  wird dabei aus einer Linearkombination von unkorrelierten Zufallsgrößen  $\mathcal{E}_t$  beziehungsweise aus benachbarten Signalen  $\mathcal{S}_k, k \neq t$  zusammengesetzt. Meist werden dabei Prozesse betrachtet, bei denen nur Zufallsgrößen aus der jüngsten Vergangenheit die Zufallsgröße  $\mathcal{S}_t$  beeinflussen. Diese Prozesse bezeichnet man als kausale finite Prozesse und kann sie grob in drei Typen einteilen:

- Moving Average Prozess der Ordnung  $q$  (MA( $q$ )-Prozess)

$$\mathcal{S}_t = \sum_{j=0}^q \beta_j \mathcal{E}_{t-j}, \quad (176)$$

- Autoregressiver Prozess der Ordnung  $p$  (AR( $p$ )-Prozess)

$$\mathcal{S}_t = \sum_{j=1}^p \alpha_j \mathcal{S}_{t-j} + \mathcal{E}_t \quad \text{und} \quad (177)$$

- Autoregressiver-Moving Average (ARMA( $p, q$ )-Prozess)

$$\mathcal{S}_t = \sum_{j=1}^p \alpha_j \mathcal{S}_{t-j} + \sum_{j=0}^q \beta_j \mathcal{E}_{t-j}, \quad (178)$$

wobei  $\mathcal{E}_t; t \in \mathbb{Z}$  unkorrelierte Zufallsvariable mit Erwartungswert  $E\{\mathcal{E}_t\} = 0$  mit einer einheitlichen Varianz  $\Sigma\{\mathcal{E}_t\} = \sigma_{\mathcal{E}}^2$  darstellen. Die zeitinvarianten (festen) Parameter  $\alpha_j, j = 1, \dots, p$  und  $\beta_j, j = 0, \dots, q$  bezeichnet man als Prozessparameter. Basierend auf dem Wold'schen Äquivalenzprinzip (Wold, 1938, Theorem 7) kann gezeigt werden, dass sich die drei Darstellungsformen für kovarianzstationäre Prozesse verlustfrei ineinander überführen lassen und somit äquivalent sind (siehe z.B. auch Scargle, 1981, Kap. IV). Die Darstellungen können sich aber sehr drastisch in der Anzahl der Parameter (wenige bis unendlich viele) unterscheiden, die zur exakten Prozessdarstellung erforderlich sind.

Im Folgenden wollen wir uns daher auf AR( $p$ )-Prozesse (177) fokussieren. Mit Hilfe der Beziehungen im *Magischen Quadrat* kann nun versucht werden, eventuell geschlossene Formeln für die vier Darstellungsformen zu entwickeln. Dies gelingt im Falle des AR( $p$ )-Prozesses und damit können nach Festlegung der Prozessparameter  $\alpha_j, j = 1, \dots, p$  alle Darstellungsformen direkt berechnet werden (siehe Abb. 6).

Für den Zeitbereich erfolgt der Übergang vom stochastischen Prozess zur Kovarianzfolge über die Bildung der Kovarianzen. Ausgehend von der Definition des autoregressiven Prozesses der Ordnung  $p$  aus (177) ergibt sich nach Multiplikation mit  $\mathcal{S}_{u-t}$  und Erwartungswertbildung die Autokorrelation  $E\{\mathcal{S}_u \mathcal{S}_{u-t}\}$  zwischen dem Signal  $\mathcal{S}_u$  und  $\mathcal{S}_{u-t}$  mit

$$\begin{aligned} E\{\mathcal{S}_u \mathcal{S}_{u-t}\} &= \alpha_1 E\{\mathcal{S}_{u-1} \mathcal{S}_{u-t}\} + \alpha_2 E\{\mathcal{S}_{u-2} \mathcal{S}_{u-t}\} + \dots \\ &\quad \dots + \alpha_p E\{\mathcal{S}_{u-p} \mathcal{S}_{u-t}\} + E\{\mathcal{E}_u \mathcal{S}_{u-t}\}. \end{aligned} \quad (179)$$

Für den Übergang auf das zweite *zentrale* Moment (Autokovarianzen) benötigen wir die Erwartungswerte  $E\{\mathcal{S}_t\}$ . Da für den Prozess  $\{\mathcal{E}_t\}_{\Delta t}$  weißes Rauschen vorausgesetzt wird, folgt nach Anwendung des Erwartungswertoperators auf (177) unmittelbar  $E\{\mathcal{S}_t\} = 0$ .

Benutzen wir die Definition  $\gamma_{|t-j|}^s := E\{\mathcal{S}_{u-j} \mathcal{S}_{u-t}\}$  und beachten  $E\{\mathcal{E}_u \mathcal{S}_u\} = \sigma_{\mathcal{E}}^2$  bzw.  $E\{\mathcal{E}_u \mathcal{S}_{u-t}\} = 0, t > 0$  so führt (179) unmittelbar auf die *Yule-Walker Gleichungen*

$$\begin{aligned} \gamma_0^s &= \alpha_1 \gamma_1^s + \alpha_2 \gamma_2^s + \dots + \alpha_p \gamma_p^s + \sigma_{\mathcal{E}}^2 \\ \gamma_1^s &= \alpha_1 \gamma_0^s + \alpha_2 \gamma_1^s + \dots + \alpha_p \gamma_{p-1}^s \\ \gamma_2^s &= \alpha_1 \gamma_1^s + \alpha_2 \gamma_0^s + \dots + \alpha_p \gamma_{p-2}^s \\ &\vdots \\ \gamma_p^s &= \alpha_1 \gamma_{p-1}^s + \alpha_2 \gamma_{p-2}^s + \dots + \alpha_p \gamma_0^s \\ \gamma_{p+1}^s &= \alpha_1 \gamma_p^s + \alpha_2 \gamma_{p-1}^s + \dots + \alpha_p \gamma_1^s \\ \gamma_{p+2}^s &= \alpha_1 \gamma_{p+1}^s + \alpha_2 \gamma_p^s + \dots + \alpha_p \gamma_2^s \\ &\vdots \end{aligned} \quad (180)$$





der den Prozess des weißen Rauschens  $\{\mathcal{E}_t\}_{\Delta t}$  durch eine Faltung im Zeitbereich beschreibt. Wegen des Faltungstheorems entspricht eine Faltung im Zeitbereich einer elementweisen Multiplikation im Frequenzbereich. Die spektrale Darstellung der Moving-Average-Koeffizienten bildet die Transferfunktion

$$H(\nu) = \sum_{j=0}^p \bar{\alpha}_j e^{-i2\pi\nu j\Delta t}, \quad \nu \in [-\nu^N, \nu^N]. \quad (186)$$

Die spektrale Repräsentation  $dZ_{\mathcal{E}}(\nu)$  des weißen Rauschens ergibt sich somit durch die Multiplikation der Transferfunktion mit der spektralen Repräsentation von  $dZ_{\mathcal{S}}(\nu)$

$$dZ_{\mathcal{E}}(\nu) = H(\nu) dZ_{\mathcal{S}}(\nu) \quad (187)$$

(vgl. Krasbutter et al., 2015). Durch einfache Umformung erhält man mit

$$dZ_{\mathcal{S}}(\nu) = \frac{dZ_{\mathcal{E}}(\nu)}{H(\nu)} = \frac{dZ_{\mathcal{E}}(\nu)}{1 - \sum_{j=1}^p \alpha_j e^{-i2\pi\nu j\Delta t}} \quad (188)$$

die spektrale Repräsentation des autoregressiven Prozesses  $\{\mathcal{S}_t\}_{\Delta t}$ . Die spektrale Verteilungsfunktion wird nach den Regeln des Magischen Quadrats durch  $\frac{1}{\Delta t} E \{dZ_{\mathcal{S}}(\nu) dZ_{\mathcal{S}}^*(\nu)\}$  gebildet. Diese Operation setzt sich zusammen aus der konjugiert komplexen Multiplikation der inversen Transferfunktion im Nenner von (188) und der konjugiert komplexen Multiplikation der spektralen Repräsentation des weißen Rauschens  $dZ_{\mathcal{E}}(\nu)$  im Zähler (188). Während der Nenner aus deterministischen Größen besteht und daher einfach ausmultipliziert werden kann, ist die Operation mit der Zufallsvariablen im Zähler weit komplexer. Da der Zähler aber lediglich die spektrale Repräsentanz des weißen Rauschens darstellt, nutzen wir zur Ableitung der spektralen Dichte/Verteilungsfunktion den unteren Übergang im Magischen Quadrat. Zunächst erfolgt der Übergang im Ortsbereich vom weißen Rauschen  $\{\mathcal{E}_t\}_{\Delta t}$  zur Kovarianzfolge  $\{\gamma_{|t|}^{\mathcal{E}}\}_{\Delta t}$ . Danach kann mit Hilfe der diskreten Form des Wiener-Chintschin Theorems (Buttkus, 2000, Kap. 8.2.2)

$$f_{\mathcal{E}} = \Delta t \sum_{t=-\infty}^{\infty} \gamma_{|t|}^{\mathcal{E}} \cos(2\pi\nu t\Delta t) \quad (189)$$

von der Kovarianzfolge  $\{\gamma_{|t|}^{\mathcal{E}}\}_{\Delta t}$  auf die spektrale Dichtefunktion  $f_{\mathcal{E}}(\nu)$  übergegangen werden (Abb. 5). Setzen wir in (189) die Kovarianzfolge für weißes Rauschen

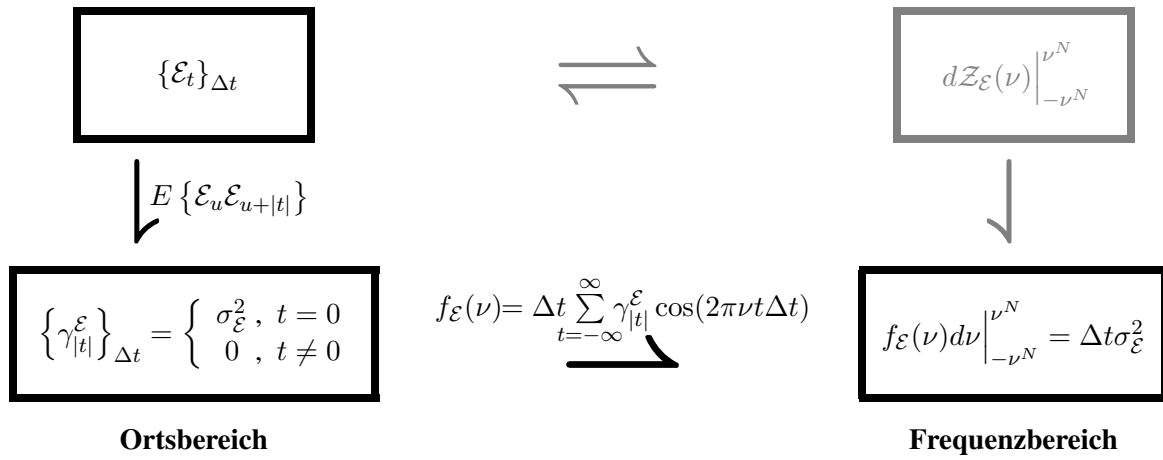
$$\gamma_{|t|}^{\mathcal{E}} = \begin{cases} \sigma_{\mathcal{E}}^2, & t = 0 \\ 0, & t \neq 0 \end{cases} \quad (190)$$

ein, so erhalten wir die spektrale Dichtefunktion des weißen Rauschens mit

$$f_{\mathcal{E}}(\nu) = \Delta t \sigma_{\mathcal{E}}^2. \quad (191)$$

Durch Zusammensetzung von Zähler und Nenner in (188) erhalten wir somit die spektrale Dichtefunktion für den autoregressiven Prozess mit

$$f_{\mathcal{S}}(\nu) = \frac{\Delta t \sigma_{\mathcal{E}}^2}{(1 + \alpha_p^2 - 2\alpha_p \cos(2\pi\nu p\Delta t) + \sum_{j=1}^{p-1} (\alpha_j^2 - 2\cos(2\pi\nu j\Delta t)(\alpha_j - \sum_{k=1}^{p-j} \alpha_k \alpha_{k+j}))} \Bigg|_{-\nu^N}^{\nu^N}. \quad (192)$$

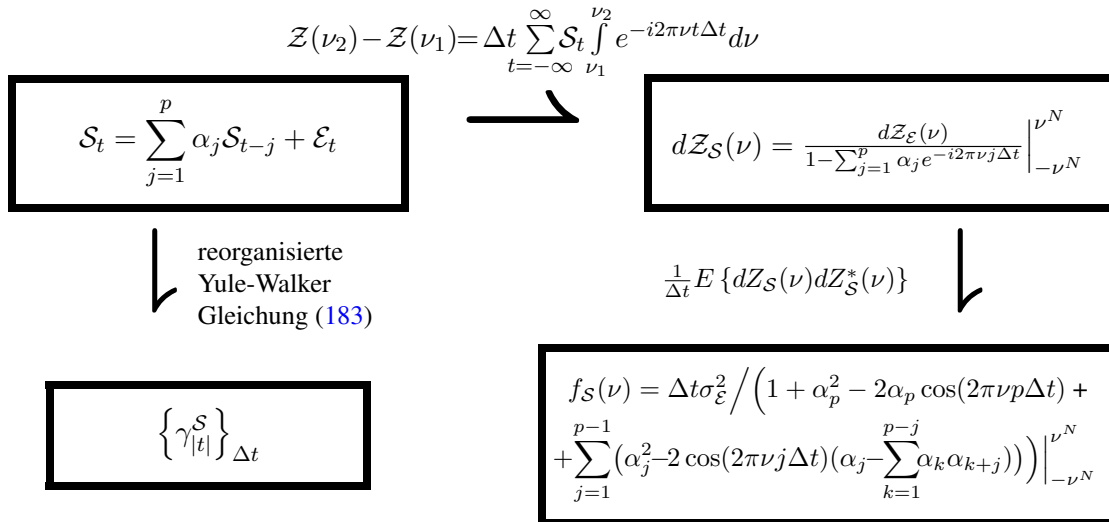


**Abb. 5:** Magisches Quadrat für weißes Rauschen. Oben links: weißes Rauschen, unten links: Autokovarianzfolge, oben rechts: spektrale Repräsentation des stochastischen Prozesses und unten rechts: spektrale Dichtefunktion (Spektraldichte). Die angeführten Formeln geben die verwendeten Übergänge in Pfeilrichtung wieder.

Im Gegensatz zur direkten Ableitung der spektralen Dichte aus der Kovarianzfunktion des autoregressiven Prozesses über das Wiener-Chintschin Theorem (Buttkus, 2000, Kap. 8.2.2)

$$f_S = \Delta t \sum_{t=-\infty}^{\infty} \gamma_{|t|}^S \cos(2\pi\nu t \Delta t) \tag{193}$$

ergibt sich bei dem hier gewählten Weg durch das Magische Quadrat über die spektrale Representation des Prozesses eine geschlossen Form mit endlichen Summen. Alle vier Darstellungsformen des autoregressiven Prozesses (177) können somit übersichtlich als Magisches Quadrat zusammengestellt werden (Abb. 6).



**Abb. 6:** Magisches Quadrat des autoregressiven Prozesses der Ordnung p

#### 4.4 Dekorrelationsfilter

Ausgangspunkt für den Einsatz des autoregressiven Prozesses zur Dekorrelation der Messungen bildet das allgemeine Kollokationsmodell (2)

$$\mathcal{L} = \mathbf{A}\boldsymbol{\xi} + \mathbf{S} + \mathcal{N}, \quad (2)$$

wobei wir zunächst ohne Einschränkung der Allgemeinheit davon ausgehen wollen, dass  $\mathbf{S}$  und  $\mathcal{N}$  den Erwartungswert Null haben,  $\mathbf{E}\{\mathbf{S}\} = \mathbf{0}_n$  und  $\mathbf{E}\{\mathcal{N}\} = \mathbf{0}_n$ . Ferner nehmen wir an, dass homogenes unkorreliertes Rauschen  $\Sigma\{\mathcal{N}\} = \mathbb{1}_n \sigma_{\mathcal{N}}^2$  vorliegt und  $\mathbf{S}$  und  $\mathcal{N}$  ebenfalls unkorreliert sind,  $\Sigma\{\mathbf{S}, \mathcal{N}\} = \mathbf{0}_n$ . Betrachten wir zunächst wieder wie in (1) die Einzelmessung  $\mathcal{L}_t$  zu einem bestimmten Zeitpunkt und modellieren wir nun das Signal  $\mathcal{S}_t$  und das Rauschen  $\mathcal{N}_t$  zum Zeitpunkt  $t$  durch einen autoregressiven Prozess der Ordnung  $p$  so erhält man

$$\mathcal{S}_t + \mathcal{N}_t = \alpha_1(\mathcal{S}_{t-1} + \mathcal{N}_{t-1}) + \alpha_2(\mathcal{S}_{t-2} + \mathcal{N}_{t-2}) + \dots + \alpha_p(\mathcal{S}_{t-p} + \mathcal{N}_{t-p}) + \mathcal{E}_t + \mathcal{N}_t, \quad (194)$$

wobei  $\mathcal{E}_t$  unkorreliertes homogenes Rauschen mit  $\mathbf{E}\{\mathcal{E}\} = \mathbf{0}_n$  und  $\Sigma\{\mathcal{E}\} = \mathbb{1}_n \sigma_{\mathcal{E}}^2$  kennzeichnet. Durch Umstellung der Gleichung (194) nach  $\mathcal{E}_t + \mathcal{N}_t$  und Anordnung der Zeitpunkte in Matrix-Vektor-Schreibweise erhält man

$$\underbrace{\begin{bmatrix} \vdots \\ \mathcal{E}_{t-2} + \mathcal{N}_{t-2} \\ \mathcal{E}_{t-1} + \mathcal{N}_{t-1} \\ \mathcal{E}_t + \mathcal{N}_t \\ \vdots \end{bmatrix}}_{\mathcal{E} + \mathcal{N}} = \underbrace{\begin{bmatrix} \ddots & & \ddots & \ddots & \ddots & & \\ & -\alpha_p & \dots & -\alpha_2 & -\alpha_1 & 1 & \\ & & -\alpha_p & \dots & -\alpha_2 & -\alpha_1 & 1 \\ & & & -\alpha_p & \dots & -\alpha_2 & -\alpha_1 & 1 \\ & & & & \ddots & & \ddots & \ddots & \ddots \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} \vdots \\ \mathcal{S}_{t-2} + \mathcal{N}_{t-2} \\ \mathcal{S}_{t-1} + \mathcal{N}_{t-1} \\ \mathcal{S}_t + \mathcal{N}_t \\ \vdots \end{bmatrix}}_{\mathbf{S} + \mathcal{N}}, \quad (195)$$

und somit eine lineare Transformation

$$\mathcal{E} + \mathcal{N} = \mathbf{H}(\mathbf{S} + \mathcal{N}) \quad (196)$$

die als ein *linearer, zeitinvarianter, nicht rekursiver, kausaler Filter* betrachtet werden kann, der den Übergang von  $\mathbf{S} + \mathcal{N}$  auf  $\mathcal{E} + \mathcal{N}$  erlaubt (Abb. 7). Der Filter bleibt unverändert mit der Zeit. Die Filtermatrix  $\mathbf{H}$  weist somit eine Toeplitz-Form auf. Da wir auf die spezielle Situation bei der Initialisierung der Zeitreihe im Moment nicht näher eingehen wollen, gehen wir zunächst von unendlichen Zeitreihen aus. Durch Varianzfortpflanzung kann der Übergang (196) auch für die Kovarianzen bzw. Kovarianzmatrizen

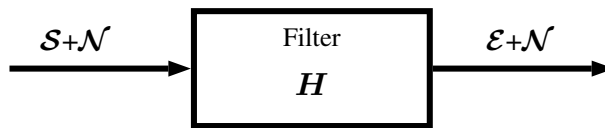


Abb. 7: Linearer Filter

hergestellt werden

$$\Sigma\{\mathcal{E}\} + \Sigma\{\mathcal{N}\} = \mathbf{H}(\Sigma\{\mathbf{S}\} + \Sigma\{\mathcal{N}\})\mathbf{H}^T \quad \text{bzw.} \quad \Sigma_{\mathcal{E}\mathcal{E}} + \Sigma_{\mathcal{N}\mathcal{N}} = \mathbf{H}(\Sigma_{\mathbf{S}\mathbf{S}} + \Sigma_{\mathcal{N}\mathcal{N}})\mathbf{H}^T, \quad (197)$$

wobei die Kovarianzmatrizen  $\Sigma_{\mathcal{E}\mathcal{E}}$  und  $\Sigma_{\mathcal{N}\mathcal{N}}$  in der Regel Diagonalmatrizen sind und nach unseren oberen Annahmen

$$\Sigma_{\mathcal{E}\mathcal{E}} + \Sigma_{\mathcal{N}\mathcal{N}} = \mathbb{1}_\infty (\sigma_{\mathcal{E}}^2 + \sigma_{\mathcal{N}}^2) \quad (198)$$

gilt. Wenden wir nun die lineare Transformation mit der Filtermatrix  $\mathbf{H}$  auf die ursprüngliche Messgleichungen (2) an, so erhalten wir das Ersatzproblem

$$\begin{aligned} \bar{\mathcal{L}} &= \bar{\mathbf{A}}\xi + \mathcal{E} + \mathcal{N} & \text{mit} & & \bar{\mathcal{L}} &= \mathbf{H}\mathcal{L}, \\ & & & & \bar{\mathbf{A}} &= \mathbf{H}\mathbf{A} \quad \text{und} \\ & & & & \mathcal{E} + \mathcal{N} &= \mathbf{H}(\mathcal{S} + \mathcal{N}), \end{aligned} \quad (199)$$

welches wegen

$$\Sigma\{\bar{\mathcal{L}}\} = \Sigma\{\mathbf{H}\mathcal{L}\} = \mathbf{H}\Sigma\{\mathcal{L}\}\mathbf{H}^T = \mathbf{H}(\Sigma_{\mathcal{S}\mathcal{S}} + \Sigma_{\mathcal{N}\mathcal{N}})\mathbf{H}^T = \Sigma_{\mathcal{E}\mathcal{E}} + \Sigma_{\mathcal{N}\mathcal{N}} \quad (200)$$

unkorrelierte Messungen aufweist. Die Parameter können somit über

$$\tilde{\mathcal{X}} = \left(\bar{\mathbf{A}}^T \bar{\mathbf{A}}\right)^{-1} \bar{\mathbf{A}}^T \bar{\mathcal{L}} \quad (201)$$

geschätzt werden. Die Varianz der geschätzten Parameter errechnet sich mit

$$\Sigma\{\tilde{\mathcal{X}}\} = \left(\bar{\mathbf{A}}^T \bar{\mathbf{A}}\right)^{-1} (\sigma_{\mathcal{E}}^2 + \sigma_{\mathcal{N}}^2). \quad (202)$$

Die Anwendung dieses Filters führt also auf dekorrelierte Zeitreihen. Die Filterung der Messungen  $\ell$  und aller Spalten der Designmatrix  $\mathbf{A}$  kann sehr effizient berechnet werden, da die Ordnung  $p$  des Filters zumeist klein ist und die Filteroperation einfach zu parallelisieren ist. Speziell bei sehr langen Zeitreihen erfolgt der Übergang von unendlichen Zeitreihen auf endliche Zeitreihen, indem die Zeitreihen nach vorne durch Nullen ergänzt werden und anschließend die ersten dekorrelierten Werte in der Parameterschätzung unberücksichtigt bleiben („Aufwärmphase“ des Filters, siehe auch Schuh (2003)). Weiterführende Verfahren und Diskussionen sind z.B. bei Schuh (1996) und Siemes (2012) zu finden. Ein Überblick über die unterschiedlichsten Verfahren zur Schätzung der AR-Parameter aus Beobachtungsreihen ist bei z.B. bei Kay und Marple (1981) zu finden.

## 5 Fazit und Ausblick

Die gewissenhafte Erfassung und Modellierung von Korrelationen in Messdaten bildet eine der Grundvoraussetzungen für eine konsistente Parameter- und Varianzschätzung. In diesem Artikel wurden einige Grundlagen zur Kollokation und Signalverarbeitung zusammengefasst, wobei speziell auf methodische Fragen eingegangen wurde. Rechentechnische Aspekte und Fragen zu Anwendungen sind sicher zu kurz gekommen, was aber den vorgegebenen Umfang gesprengt hätte. Hier mögen die entsprechenden Literaturhinweise weiterhelfen. Weiterführende Forschung ist jedoch erforderlich um speziell zeitvariable multivariate stochastische Prozesse besser zu verstehen und vielfach ungerechtfertigte Annahmen wie Homogenität und Isotropie zu überwinden.

## Literatur

- Bochner, S. (2005): *Harmonic Analysis and the Theory of Probability*. Dover Books on Mathematics. Dover Publication Inc. 24, 25
- Bottoni, G., R. Barzaghi (1993): Fast collocation. *Bulletin Géodésique*, 67:119–126. 26
- Brockwell, P. J., R. A. Davis (1991): *Time Series: Theory and Methods*. Second Edition. Springer, Berlin. ISBN 1-4419-0319-8. 3, 4, 24
- Buttkus, B. (2000): *Spectral Analysis and Filter Theory in Applied Geophysics*. Springer-Verlag, Berlin, Heidelberg. 5, 33, 34
- Cressie, N. A. (1991): *Statistics for Spatial Data*. Wiley. 6
- Eren, K. (1982): Toeplitz matrices and frequency domain collocation. *Manuscripta Geodaetica*, 7:85–118. 26

- Gaspari, G., S. Cohn (1999): Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757. 26, 27
- Gaspari, G., S. Cohn, J. Guo, S. Pawson (2006): Construction and application of covariance functions with variable length-fields. *Quarterly Journal of the Royal Meteorological Society*, 132:1815–1838. 27
- Grafarend, E. W. (1976): Geodetic applications of stochastic processes. *Physics of the Earth and Planetary Interiors*, 12(2-3):151 – 179. ISSN 0031-9201. doi:10.1016/0031-9201(76)90045-5. 6
- Heiskanen, W., H. Moritz (1967): *Physical Geodesy*. Institute of Physical Geodesy, Technical University, Graz. Corrected reprint of the original edition by W.H. Freeman and Company, San Franzisco, 1967. 26
- Hirvonen, R. (1962): On the statistical analysis of gravity anomalies. *Reports of the Department of Geodetic Science*, Band 19. Ohio State University (OSU), Ohio. 26
- Jäggi, A. (2007): Pseudo-stochastic orbit modelling of Low Earth Satellites using the Global Positioning System. *Geodätisch-geophysikalische Arbeiten in der Schweiz*, Band 73. Schweizerische Geodätische Kommission. URL <http://boris.unibe.ch/id/eprint/25278>. 10
- Kay, S., S. Marple (1981): Spectrum Analysis - A Modern Perspective. *Proceedings of the IEEE*, 69(11):1380–1419. ISSN 0018-9219. 36
- Khintchine(Chintschin), A. (1934): Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109:604 – 615. URL [http://resolver.sub.uni-goettingen.de/purl?PPN235181684\\_0109](http://resolver.sub.uni-goettingen.de/purl?PPN235181684_0109). 3
- Koch, K., H. Kuhlmann, W.-D. Schuh (2010): Approximating covariance matrices estimated in multivariate models by estimated auto- and cross-covariances. *J. Geodesy*, 84(6):383–397. doi:10.1007/s00190-010-0375-5. 28, 30
- Kolmogorov, A. N. (1941): Interpolation and extrapolation in stationary random sequences (in Russian). *Izv. Akad. Nauk SSSR Ser. Mat.*, 5:3–14. (English: Selected Works of A.N. Kolmogorov, Volume II: Probability Theory and Mathematical Statistics. Editor: Shirayev A.N. Series: Mathematics and Its Applications, Vol. 26, 1992, Kluwer Academic Publishers, p. 272-280). 6
- Krarp, T. (1969): A contribution to the mathematical foundation of physical geodesy. *Geodätisches Institut, Meddelelse n. 44*. 1, 6
- Krasbutter, I., B. Kargoll, W.-D. Schuh (2015): Magic Square of Real Spectral and Time Series Analysis with an Application to Moving Average Processes. IAG Symposia 140 S. 9–14. Springer. doi:10.1007/978-3-319-10828-5\_2. 5, 33
- Kreyszig, E. (1993): *Advanced engineering mathematics*. John Wiley & Sons, 7. Auflage. 25, 28
- Matheron, G. (1963): Principles of geostatistics. *Economic geology*, 58(8):1246–1266. ISSN 0361-0128 (P), 1554-0774 (E). 6
- Mayer-Gürr, T. (2008): Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE. *Schriftenreihe des Instituts für Geodäsie und Geoinformation*, Band 9. Universität Bonn. URL <http://nbn-resolving.de/urn:nbn:de:hbz:5N-09047>. 10
- Meissl, P. (1971): A study of covariance functions related to the earth's disturbing potential. *Reports of the Department of Geodetic Science*, Band 151. Ohio State University (OSU), Ohio. 26
- Meissl, P. (1982): Least squares adjustment: A modern approach. *Mitteilungen der Geodätischen Institute der TU Graz*, Band 43. Geodätischen Institute der TU Graz, Graz. URL [ftp://skylab.itg.uni-bonn.de/schuh/Separata\\_Meissl/meissl\\_1982b.pdf](ftp://skylab.itg.uni-bonn.de/schuh/Separata_Meissl/meissl_1982b.pdf). 8, 12
- Moreaux, G. (2008): Compactly supported radial covariance functions. *Journal of Geodesy*, 82(7):431–443. ISSN 0949-7714. doi:10.1007/s00190-007-0195-4. 27
- Moritz, H. (1972): Advanced Least-Squares Methods. *Reports of the Department of Geodetic Science*, Band 175. Ohio State University (OSU), Ohio. 1, 6
- Moritz, H. (1973): *Least-squares collocation*. Deutsche Geodätische Kommission, München. Reihe A, 75. 7, 30
- Moritz, H. (1976): Covariance functions in least-squares collocation. *Reports of the Department of Geodetic Science*, Band 240. Ohio State University (OSU), Ohio. 26
- Moritz, H. (1980): *Advanced Physical Geodesy*. Wichmann, Karlsruhe. 6, 7, 8, 9, 10, 13, 26, 30
- Mussio, L. (1984): Il metodo della collocazione minimi quadrati e le sue applicazioni per l'analisi statistica dei risultati delle compensazioni. Cunietti, M., (Hrsg.), *Ricerche di Geodesia, Topografia e Fotogrammetria*, 4, S. 305–338. CLUP, Milano. 26
- Papoulis, A. (1977): *Signal Analysis*. International Student Edition. McGraw-Hill. 27, 29
- Reguzzoni, M., F. Sansò, G. Venuti (2005): The theory of general kriging, with applications to the determination of a local geoid. *Geophysical Journal*, 162(2):303–314. doi:10.1111/j.1365-246X.2005.02662.x. 7
- Rummel, R. (1975): Zur Behandlung von Zufallsfunktionen und -folgen in der physikalischen Geodäsie. DGK, Reihe C 208, Verlag der Bayerischen Akademie der Wissenschaften. 3
- Rummel, R. (1976): A model comparison in least squares collocation. *Bulletin Géodésique*, 50(2):181–192. ISSN 0007-4632. doi:10.1007/BF02522317. URL <http://dx.doi.org/10.1007/BF02522317>. 6
- Sansò, F. (1986): Statistical methods in physical geodesy. Sünkel, H., (Hrsg.), *Mathematical and Numerical Techniques in Physical Geodesy*, Band 7 *Lecture Notes in Earth Sciences*, S. 49–155. Springer Berlin Heidelberg. ISBN 978-3-540-16809-6. doi:10.1007/BFb0010132. 6
- Sansò, F., W.-D. Schuh (1987): Finite covariance functions. *Bulletin Géodésique*, 61:331–347. 26, 27

- Sansò, F., C. Tscherning (2003): Fast spherical collocation: theory and examples. *J. Geodesy*, 77:101–112. 26
- Scargle, J. (1981): Studies in astronomical time series analysis. I – Modeling random processes in the time domain. *Astrophysical Journal Supplement Series*, 45:1–71. URL <http://articles.adsabs.harvard.edu/full/1981ApJS...45....1S>. 31
- Schuh, W.-D. (1989): Kollokation - zu rechenaufwendig? *ZAMM, Z. angew. Math. Mech.*, 69, 4, T73-T75. URL <http://onlinelibrary.wiley.com/doi/10.1002/zamm.19890690403/pdf>. 28
- Schuh, W.-D. (1996): Tailored numerical solution strategies for the global determination of the earth's gravity field. *Mitteilungen der Technischen Universität Graz, Graz*. Folge 81. 26, 36
- Schuh, W.-D. (2003): The processing of band-limited measurements; filtering techniques in the least squares context and in the presence of data gaps. *Space Science Reviews*, 108(1-2):67–78. doi:10.1023/A:1026121814042. 36
- Schuh, W.-D., I. Krasbutter, B. Kargoll (2014): Korrelierte Messung - was nun? Neuner, H., (Hrsg.), *Zeitabhängige Messgrößen - Ihre Daten haben (Mehr-)Wert*, Band 74 *DVW-Schriftenreihe*, S. 85 – 101. Wißner, Augsburg. ISBN: 978-3-89639-970-0. 32
- Schwarz, K. (1976): Least squares collocation for large systems. *Bollettino di Geodesia e Scienze Affini*, No. 3:309–324. Presented at “Sixth Symposium on Mathematical Geodesy”, Siena (April 2-5, 1975). 6
- Siemes, C. (2012): Digital Filtering Algorithms for Decorrelation within Large Least Squares Problems. *Schriftenreihe des Instituts für Geodäsie und Geoinformation*, Band 32. Rheinischen Friedrich-Wilhelms-Universität Bonn, Bonn. URL <http://nbn-resolving.de/urn:nbn:de:hbz:5N-13749>. 36
- Stewart, J. (1976): Positive definite functions and generalizations, a historical survey. *Journal of Rocky Mountains Mathematics*, 6:409–434. 25
- Tscherning, C. (1972): Representation of covariance functions related to the anomalous potential of the earth using reproducing kernels. Internal Report. Danish Geodetic Institute, Kobenhavn. No. 3. 26
- Tscherning, C., P. Knudsen, R. Forsberg (1994): Description of the GRAVSOFTE package. Technischer Report, Geophysical Institute, University of Copenhagen. 26
- Tscherning, C., R. Rapp (1974): Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models. *Reports of the Department of Geodetic Science*, Band 208. Ohio State University (OSU), Ohio. 26
- Wackernagel, H. (2003): *Multivariate geostatistics: An Introduction with Applications*. Springer, 3. Auflage. 6
- Wiener, N. (1949): *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. John Wiley Sons, New York. 6
- Wold, H. (1938): *A study in the analysis of stationary time series*. Almqvist & Wiksell, Stockholm, Univ. Diss., 2. Auflage. 3, 31
- Wolf, H. (1974): Über verallgemeinerte Kollokation. *Zeitschrift für Vermessungswesen*, 99:475–478. 6