

## Modele zapisane w przestrzeni stanów

- Modele Przestrzeni Stanów (*State Space Models*) są to modele, w których część parametrów jest nieobserwowalna i losowa.
- Zachowanie wielowymiarowej zmiennej  $y_t$  zależy więc od
  - parametrów deterministycznych
  - wektora zmiennych egzogenicznych  $X_t$
  - nieobserwowalnego stanu danego wektorem  $z_t$ .
- Równanie pomiarowe (*measurement equation*)

$$y_t = H_t z_t + G_t x_t + v_t$$

gdzie  $z_t$  nazywamy wektorem stanu (*state vector*)

- Równanie przejścia (*transition equation*)

$$\mathbf{z}_t = \mathbf{B}_t \mathbf{z}_{t-1} + \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t$$

- Razem oba te równania nazywamy liniowym systemem przestrzeni stanów (*linear state space system*) gdzie
- $\mathbf{y}_t$  jest  $(G \times 1)$  wymiarowym obserwowalnym wektorem zmiennych endogenicznych
- $\mathbf{z}_t$  jest  $(S \times 1)$  wymiarowym wektorem stanów natury
- $\mathbf{x}_t$  jest  $(K \times 1)$  wymiarowym obserwowalnym wektorem zmiennych wejściowych
- $\mathbf{v}_t$  jest  $(G \times 1)$  wymiarowym błędów pomiaru

- $w_t$  jest  $(S \times 1)$  wymiarowym błędów losowych w równaniach przejścia
- $H_t$  jest  $(G \times S)$  wymiarową macierzą pomiarową
- $G_t$  jest  $(G \times K)$  wymiarową macierzą wejścia w równaniach pomiaru
- $B_t$  jest  $(S \times S)$  wymiarową macierzą przejścia
- $F_t$  jest  $(S \times K)$  wymiarową macierzą wejścia w równaniach przejścia
- O macierzach  $H_t, G_t, B_t, F_t$  zakładamy, że są niezależne od  $v_t$ .
- Zakładamy też, że proces zaczyna się od pewnego stanu początkowego  $z_0, x_0$ .
- Zaburzenia losowe  $w_t, v_t$  są niezależne i mają stałe w czasie wariancje

- łączny proces generujący zaburzenia losowe jest niekorelowany w czasie, ma rozkład

$$\begin{bmatrix} \boldsymbol{w}_t \\ \boldsymbol{v}_t \end{bmatrix} \sim N \left( 0, \begin{bmatrix} \boldsymbol{\Sigma}_w & 0 \\ 0 & \boldsymbol{\Sigma}_v \end{bmatrix} \right) \text{ dla } t = 0, \dots$$

i jest niezależny od  $\boldsymbol{z}_0$ , który ma rozkład:

$$\boldsymbol{z}_0 \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

- Przyjmijmy następujące oznaczenia

$$\mathbf{Y}_s = (\mathbf{y}_1, \dots, \mathbf{y}_s)$$

$$\mathbf{z}_{t|s} = \mathbb{E}(\mathbf{z}_t | \mathbf{Y}_s)$$

$$\Sigma_z(t|s) = \text{Var}(\mathbf{z}_t | \mathbf{Y}_s)$$

$$\mathbf{y}_{t|s} = \mathbb{E}(\mathbf{y}_t | \mathbf{Y}_s)$$

$$\Sigma_y(t|s) = \text{Var}(\mathbf{y}_t | \mathbf{Y}_s)$$

$$\Sigma_{zy}(t|s) = \text{Cov}(\mathbf{z}_t, \mathbf{y}_t | \mathbf{Y}_s)$$

- $(\mathbf{z} | \mathbf{y}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  oznacza, że warunkowy rozkład  $\mathbf{z}$  przy danym  $\mathbf{y}$  jest wielowymiarowym rozkładem normalnym o wartości oczekiwanej  $\boldsymbol{\mu}$  i macierzy wariancji kowariancji  $\boldsymbol{\Sigma}$ .
- Przy tych założeniach warunkowe rozkłady  $\mathbf{z}_t$  i  $\mathbf{y}_t$  będą rozkładami normalnymi o następujących parametrach

$$(\mathbf{z}_t | \mathbf{Y}_{t-1}) \sim N(\mathbf{z}_{t|t-1}, \boldsymbol{\Sigma}_z(t|t-1)) \text{ dla } t = 2, \dots, T$$

$$(\mathbf{z}_t | \mathbf{Y}_t) \sim N(\mathbf{z}_{t|t}, \boldsymbol{\Sigma}_z(t|t)) \text{ dla } t = 1, \dots, T$$

$$(\mathbf{y}_t | \mathbf{Y}_{t-1}) \sim N(\mathbf{y}_{t|t-1}, \boldsymbol{\Sigma}_y(t|t-1)) \text{ dla } t = 2, \dots, T$$

$$(\mathbf{z}_t | \mathbf{Y}_T) \sim N(\mathbf{z}_{t|T}, \boldsymbol{\Sigma}_z(t|T)) \text{ dla } t > T$$

$$(\mathbf{y}_t | \mathbf{Y}_T) \sim N(\mathbf{y}_{t|T}, \boldsymbol{\Sigma}_y(t|T)) \text{ dla } t > T$$

## Przykład

- Proces  $MA(2)$

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

- równanie pomiarowe

$$y_t = \mathbf{H} \mathbf{z}_t$$

$$\mathbf{H} = [ 1 \quad \theta_1 ]$$

$$\mathbf{z}'_t = [ \varepsilon_t \quad \varepsilon_{t-1} ]$$

- równanie przejścia

$$\mathbf{z}_t = \mathbf{B} \mathbf{z}_{t-1} + \mathbf{w}_t$$
$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{w}_t = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}, \quad \Sigma_w = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & 0 \end{bmatrix}$$

# Filtry Kalmana

- Przy tych założeniach można dowieść następujących zależności nazywanych rekursjami filtrów Kalmana (*Kalman filter recursions*)

**Inicjacja:** ( $t = 0$ )

$$\begin{aligned}z_{0|0} &= \mu_0 \\ \Sigma_z(0|0) &= \Sigma_0\end{aligned}$$



**Predykcja:**  $(1 \leq t \leq T)$

$$\begin{aligned}z_{t|t-1} &= \mathbf{B}_t z_{t-1|t-1} + \mathbf{F}_t \mathbf{x}_{t-1} \\ \Sigma_z(t|t-1) &= \mathbf{B}_t \Sigma_z(t-1|t-1) \mathbf{B}'_t + \Sigma_w \\ \mathbf{y}_{t|t-1} &= \mathbf{H}_t z_{t|t-1} + \mathbf{G}_t \mathbf{x}_t \\ \Sigma_y(t|t-1) &= \mathbf{H}_t \Sigma_z(t|t-1) \mathbf{H}'_t + \Sigma_v\end{aligned}$$

**Korekta:**  $(1 \leq t \leq T)$

$$\begin{aligned}z_{t|t} &= z_{t|t-1} + \mathbf{P}_t \left( \mathbf{y}_t - \mathbf{y}_{t|t-1} \right) \\ \Sigma_z(t|t) &= \Sigma_z(t|t-1) - \mathbf{P}_t \Sigma_y(t|t-1) \mathbf{P}'_t\end{aligned}$$

gdzie

$$P_t = \Sigma_z(t|t-1) \mathbf{H}'_t \Sigma_y(t|t-1)^{-1} \quad ((\text{zysk z filtru Kalmana}))$$

- W przypadku, kiedy nie istnieje odwrotność macierzy  $\Sigma_y(t|t-1)$  możemy zastosować uogólnioną odwrotność macierzy.
- Rekursje przeprowadzamy zaczynając od predykcji dla  $t = 1$ .
- Potem przeprowadzamy korektę dla  $t = 1$ .
- Później odpowiednio predykcję i korektę przeprowadzamy dla  $t = 2, t = 3$  itd.

**Prognozowanie:** ( $t > T$ )

$$\begin{aligned}
 z_{t|T} &= \mathbf{B}_t z_{t-1|T} + \mathbf{F}_t \mathbf{x}_{t-1} \\
 \Sigma_z(t|T) &= \mathbf{B}_t \Sigma_z(t-1|T) \mathbf{B}'_t + \Sigma_w \\
 \mathbf{y}_{t|T} &= \mathbf{H}_t z_{t|T} + \mathbf{G}_t \mathbf{x}_t \\
 \Sigma_y(t|T) &= \mathbf{H}_t \Sigma_z(t|T) \mathbf{H}'_t + \Sigma_v
 \end{aligned}$$

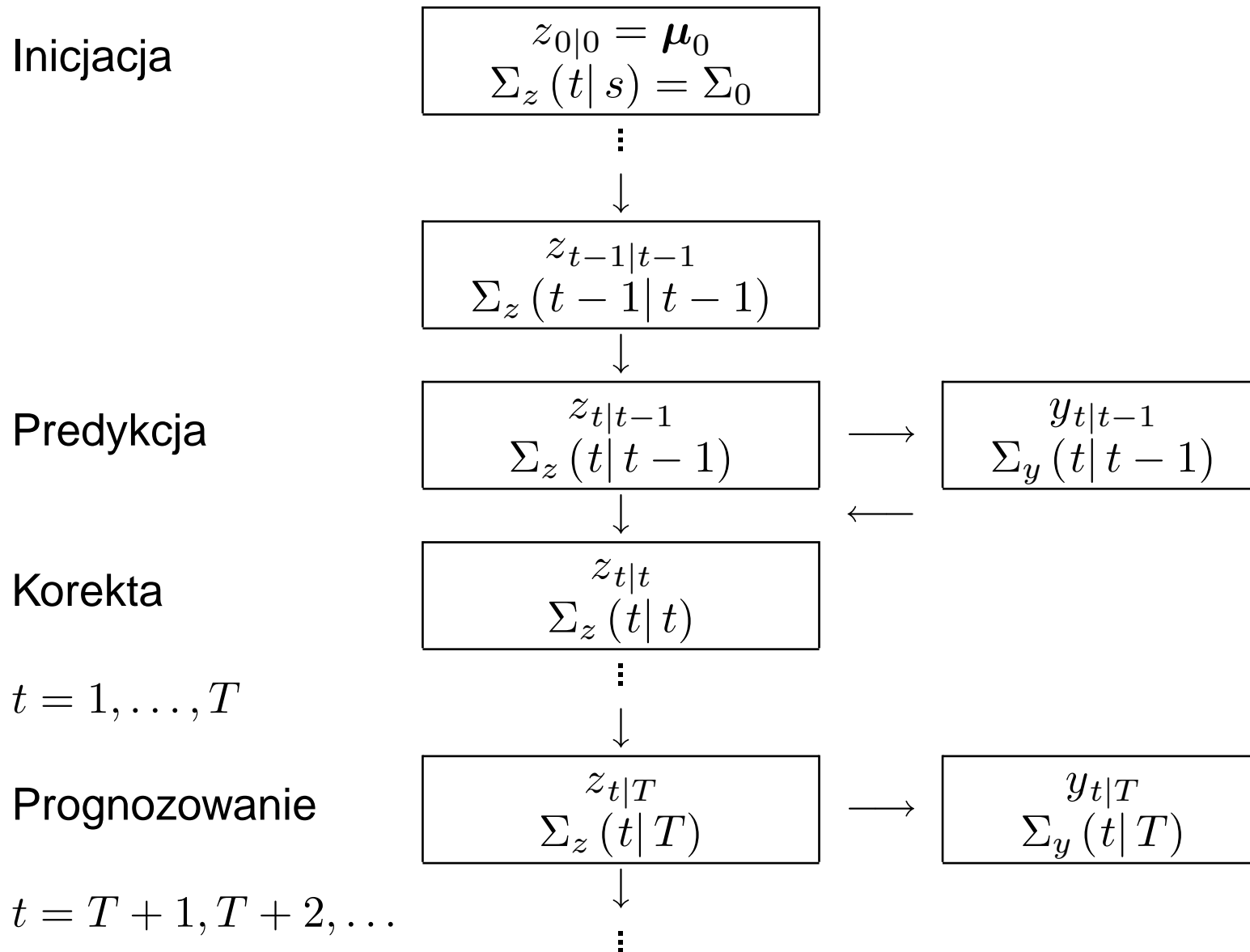
- Prognozowanie wykonujemy rekursywnie dla  $t = T + 1, T + 2, \dots$
- Niekiedy chcemy oszacować wartości wektora stanu dla znanego  $\mathbf{Y}_T$ .
- Posługujemy się do tego wzorami rekursywnymi nazywanymi wygładzaniem Kalmana.
- Wzory te stosujemy kolejno dla  $T - 1, T - 2, \dots$

**Wygładzanie: ( $t < T$ )**

$$z_{t|T} = z_{t|t} + \mathbf{S}_t (z_{t+1|T} - z_{t+1|t})$$
$$\Sigma_z (t|T) = \Sigma_z (t|t) - \mathbf{S}_t [\Sigma_z (t+1|t) - \Sigma_z (t+1|T)] \mathbf{S}'_t$$

gdzie

$$\mathbf{S}_t = \Sigma_z (t|t) \mathbf{B}'_{t+1} \Sigma_z (t+1|t)^{-1} \quad ((\text{macierz wygładzania Kalmana}))$$



## Przykład algorytm w zastosowaniach

- Proces  $MA(2)$

- Inicjacja (załóżmy, że  $\varepsilon_{0|0}$  i  $\varepsilon_{-1|0}$  są niezależne i mają równe wariancje)

$$\mathbf{z}_{0|0} = \begin{bmatrix} \varepsilon_{0|0} & \varepsilon_{-1|0} \end{bmatrix}$$
$$\boldsymbol{\Sigma}_z(0|0) = \sigma_{11}(0|0) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

– Predykcja:

$$\begin{aligned}z'_{t|t-1} &= \begin{bmatrix} 0 & \varepsilon_{t-1|t-1} \end{bmatrix} \\ \Sigma_z(t|t-1) &= \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_{11}(t-1|t-1) \end{bmatrix} \\ \mathbf{y}_{t|t-1} &= \boldsymbol{\theta}_1 \varepsilon_{t-1|t-1} \\ \Sigma_y(t|t-1) &= \sigma_w^2 + \boldsymbol{\theta}_1^2 \sigma_{11}(t-1|t-1)\end{aligned}$$

– Korekta:

$$P_t = \left[ \sigma_w^2 + \boldsymbol{\theta}_1 \sigma_{11}(t-1|t-1) \right]^{-1} \begin{bmatrix} \sigma_w^2 \\ \boldsymbol{\theta}_1 \sigma_{11}(t-1|t-1) \end{bmatrix}$$

$$z_{t|t} = z_{t|t-1} + P_t (y_t - y_{t|t-1})$$

$$\Sigma_z (t|t) = \Sigma_z (t|t-1) - P_t \Sigma_y (t|t-1) P_t'$$

– Prognozowanie  $t > T$

$$z'_{t|T} = \begin{bmatrix} 0 & \varepsilon_{t-1|T} \end{bmatrix}$$

$$\Sigma_z (t|T) = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_{11} (t-1|T) \end{bmatrix}$$

$$y_{t|T} = \theta_1 \varepsilon_{t-1|T}$$

$$\Sigma_y (t|T) = \sigma_w^2 + \theta_1 \sigma_{11} (t-1|T)$$



– Wygładzanie  $t < T$

$$\mathbf{S}_t = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$\mathbf{z}_{t|T} = \begin{bmatrix} \varepsilon_{t|t} + (\varepsilon_{t+1|T} - \varepsilon_{t+1|t}) \\ \varepsilon_{t-1|t} \end{bmatrix}$$

# Estymacja parametrów w Modelach Przestrzeni Stanów

- Jedną z podstawowych zalet modeli przestrzeni stanów jest to, że umożliwiają one stosunkowo łatwą specyfikację całej klasy modeli dla analizy szeregów czasowych.
- Zgromadźmy wszystkie nieznanne parametry z macierzy  $B_t$ ,  $F_t$ ,  $H_t$ ,  $\Sigma_w$ ,  $\Sigma_v$ ,  $\Sigma_0$  i  $\mu_0$  w wektorze parametrów  $\delta$ .
- Estymacja parametrów modeli przestrzeni stanów polegać będzie na maksymalizacji funkcji wiarygodności względem parametru  $\delta$ .
- Okazuje się, że postać funkcji wiarygodności można stosunkowo prosto przestawić przy zastosowaniu Filtrów Kalmana.

## Funkcja wiarygodności

- Korzystając z twierdzenia Bayesa można przedstawić łączną funkcję wiarygodności dla zależnych zdarzeń  $\mathbf{Y}_T = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T)$  w postaci następującego iloczynu

$$\begin{aligned} L(\mathbf{Y}_T, \boldsymbol{\theta}) &= L(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T, \boldsymbol{\theta}) = f(\mathbf{y}_1, \boldsymbol{\theta}) f(\mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{y}_1, \boldsymbol{\theta}) \\ &= f(\mathbf{y}_1, \boldsymbol{\theta}) f(\mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{Y}_1, \boldsymbol{\theta}) \\ &= f(\mathbf{y}_1, \boldsymbol{\theta}) f(\mathbf{y}_2 | \mathbf{Y}_1, \boldsymbol{\theta}) f(\mathbf{y}_3, \dots, \mathbf{y}_T | \mathbf{Y}_2, \boldsymbol{\theta}) \\ &\vdots \\ &= f(\mathbf{y}_1, \boldsymbol{\theta}) f(\mathbf{y}_2 | \mathbf{Y}_1, \boldsymbol{\theta}) \times \dots \times f(\mathbf{y}_T | \mathbf{Y}_{T-1}, \boldsymbol{\theta}). \end{aligned}$$

- Logarytm funkcji wiarygodności będzie miał więc postać

$$\ell(\mathbf{Y}_T, \boldsymbol{\theta}) = \ln f(\mathbf{y}_1, \boldsymbol{\theta}) + \sum_{i=1}^T \ln f(\mathbf{y}_t | \mathbf{Y}_{t-1}, \boldsymbol{\theta})$$

- Jeśli prawdziwe są założenie konieczne do wyprowadzenia postaci Filtrów Kalmana to zmienna  $(\mathbf{y}_t | \mathbf{Y}_{t-1}, \boldsymbol{\theta})$  będzie miała rozkład normalny a funkcja wiarygodności przyjmie postać

$$\begin{aligned} \ell(\mathbf{Y}_T, \boldsymbol{\theta}) = & -\frac{KT}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^T \ln |\boldsymbol{\Sigma}_y(t|t-1)| \\ & - \sum_{t=1}^T \left( \mathbf{y}_t - \mathbf{y}_{t|t-1} \right)' \boldsymbol{\Sigma}_y^{-1}(t|t-1) \left( \mathbf{y}_t - \mathbf{y}_{t|t-1} \right). \end{aligned}$$

- Wartość warunkowej wariancji  $\Sigma_y(t|t-1)$  i warunkowej wartości oczekiwanej  $\mathbf{y}_{t|t-1}$  otrzymujemy rekurencyjnie ze wzorów na Filtry Kalmana.
- Dla uproszczenia notacji oznaczmy

$$\mathbf{e}_t(\boldsymbol{\theta}) = \mathbf{y}_t - \mathbf{y}_{t|t-1}$$

i

$$\Sigma_t(\boldsymbol{\theta}) = \Sigma_y^{-1}(t|t-1)$$

- Przy takim zapisie funkcja wiarygodności przyjmuje postać

$$\ell(\mathbf{Y}_T, \boldsymbol{\theta}) = -\frac{KT}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^T [\ln |\Sigma_t(\boldsymbol{\theta})| + \mathbf{e}'_t(\boldsymbol{\theta}) \Sigma_t^{-1}(\boldsymbol{\theta}) \mathbf{e}_t]$$

- Do szacowanych parametrów modelu należeć mogą dowolne parametry będące częścią macierzy  $H_t, G_t, B_t, F_t$  oraz macierze  $\Sigma_w, \Sigma_v$  oraz  $\mu_0, \Sigma_0$ .
- W wielu zastosowaniach okazuje się, że trzeba na parametry nałożyć dodatkowe ograniczenia celem uzyskania identyfikacji.

## Własności estymatorów $MNW$

- W przypadku filtrów Kalmana nie jest łatwo podać ogólne warunki, dla których estymatory  $MNW$  mają standardowe własności.
- Przy spełnionych warunkach z początku tego rozdziału najłatwiej podać takie warunki dla przypadku, gdy macierz  $G_t = G$ ,  $B_t = B$ ,  $F_t = F$  nie zależą od  $t$ . W takim przypadku, jeśli
  1. wektor  $\theta$  znajduje się we wnętrzu zbioru  $\Theta$
  2.  $H_t = (x_t \otimes I) J$ , gdzie  $J$  jest z góry znaną macierzą lub  $H_t = H$  jest nielosową macierzą parametrów

3.  $\mathbf{x}_t$  jest niestochastyczne i istnieją takie stałe  $c_1$  i  $c_2$ , że  $c_1 \leq \mathbf{x}'_t \mathbf{x}_t \leq c_2$  dla  $t = 1, 2, \dots$
4. spełniony jest warunek, że macierz informacyjna jest asymptotycznie nieosobliwa
5. wszystkie wartości własne macierzy  $B$  są co do modułu mniejsze od 1

to estymator

$$\sqrt{n} (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \mathbf{i}^{-1}(\boldsymbol{\theta}))$$