Derivation of Kalman Prediction Step

Recall prediction step of sequential Bayesian estimation

$$\underbrace{p(\boldsymbol{x}|\boldsymbol{y}_-)}_{\substack{\text{Predicted}\\\text{posterior pdf}}} = \int p(\boldsymbol{x}|\boldsymbol{x}_-)\underbrace{p(\boldsymbol{x}_-|\boldsymbol{y}_-)}_{\substack{\text{Previous}\\\text{posterior pdf}}}\mathrm{d}\boldsymbol{x}_-$$

$$= \int p(\boldsymbol{x}|\boldsymbol{x}_-,\boldsymbol{y}_-)p(\boldsymbol{x}_-|\boldsymbol{y}_-)\mathrm{d}\boldsymbol{x}_- \qquad \text{(Markov Assumptions)}$$

$$= \int p(\boldsymbol{x},\boldsymbol{x}_-|\boldsymbol{y}_-)\mathrm{d}\boldsymbol{x}_-$$

Notation:

$$\boldsymbol{x} = \boldsymbol{x}_n$$
$$\boldsymbol{y} = \boldsymbol{y}_n$$
$$\boldsymbol{u} = \boldsymbol{u}_n$$
$$\boldsymbol{x}_- = \boldsymbol{x}_{n-1}$$
$$\boldsymbol{y}_- = \boldsymbol{y}_{1:n-1}$$
$$\boldsymbol{G} = \boldsymbol{G}_n$$

Note that $p(\boldsymbol{x}_-|\boldsymbol{y}_-) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}, \boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-})$

Part I: First, we show that, given $\boldsymbol{y}_-$, $\boldsymbol{w} = [\boldsymbol{x}_-^{\mathrm{T}}, \boldsymbol{x}^{\mathrm{T}}]^{\mathrm{T}}$ follows a joint Gaussian distribution, i.e.,

$$p(\boldsymbol{x},\boldsymbol{x}_-|\boldsymbol{y}_-) = p(\boldsymbol{w}|\boldsymbol{y}_-) \propto \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{w}}, \boldsymbol{\Sigma}_{\boldsymbol{ww}})$$

with mean and covariance given by

$$\boldsymbol{\mu}_{\boldsymbol{w}} = \begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\\ \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\\ \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\end{bmatrix} \qquad \boldsymbol{\Sigma}_{\boldsymbol{ww}} = \begin{bmatrix}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-} & \boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}|\boldsymbol{y}_-}\\ \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}_-|\boldsymbol{y}_-} & \boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}_-}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-} & \boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{G}^{\mathrm{T}}\\ \boldsymbol{G}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-} & \boldsymbol{G}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{G}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{uu}}\end{bmatrix}$$

1

Part I (cnt.): Note that $p(\boldsymbol{x}|\boldsymbol{x}_-,\boldsymbol{y}_-) = p(\boldsymbol{x}|\boldsymbol{x}_-) = \mathcal{N}(\boldsymbol{G}\boldsymbol{x}_-, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ and $p(\boldsymbol{x}_-|\boldsymbol{y}_-) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}, \boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-})$ are both Gaussian distributed. Thus, we have that $p(\boldsymbol{w}|\boldsymbol{y}_-) = p(\boldsymbol{x},\boldsymbol{x}_-|\boldsymbol{y}_-) = p(\boldsymbol{x}|\boldsymbol{x}_-,\boldsymbol{y}_-)p(\boldsymbol{x}_-|\boldsymbol{y}_-)$ follows a joint Gaussian distribution (as derived in a previous class; see slides 5). The missing parameters of the joint Gaussian distribution can be calculated as follows

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-} = \mathrm{E}(\boldsymbol{G}\boldsymbol{x}_- + \boldsymbol{u}) = \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}$$

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}_-} &= \mathrm{E}\big[(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-})(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-})^{\mathrm{T}}\big]\\ &= \mathrm{E}\big[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\big] - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-}\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-}^{\mathrm{T}}\\ &= \mathrm{E}\big[(\boldsymbol{G}\boldsymbol{x}_- + \boldsymbol{u})(\boldsymbol{G}\boldsymbol{x}_- + \boldsymbol{u})^{\mathrm{T}}\big] - \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}}\\ &= \boldsymbol{G}\mathrm{E}[\boldsymbol{x}_-\boldsymbol{x}_-^{\mathrm{T}}]\boldsymbol{G}^{\mathrm{T}} + \mathrm{E}\big[\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\big] - \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}}\\ &= \boldsymbol{G}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{G}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{uu}}\end{aligned}$$

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}_-|\boldsymbol{y}_-} &= \mathrm{E}\big[(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-})(\boldsymbol{x}_- - \boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-})^{\mathrm{T}}\big]\\ &= \boldsymbol{G}\mathrm{E}[\boldsymbol{x}_-\boldsymbol{x}_-^{\mathrm{T}}] - \boldsymbol{G}\mathrm{E}[\boldsymbol{x}_-]\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}}\\ &= \boldsymbol{G}(\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-} + \boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}})\\ &\qquad - \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}}\\ &= \boldsymbol{G}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-}\end{aligned}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}|\boldsymbol{y}_-} = \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}_-|\boldsymbol{y}_-}^{\mathrm{T}}$$

$$\mathrm{E}[\boldsymbol{x}_-\boldsymbol{x}_-^{\mathrm{T}}] = \boldsymbol{\Sigma}_{\boldsymbol{x}-\boldsymbol{x}-|\boldsymbol{y}-} + \boldsymbol{\mu}_{\boldsymbol{x}-|\boldsymbol{y}-}\boldsymbol{\mu}_{\boldsymbol{x}-|\boldsymbol{y}-}^{\mathrm{T}}$$

Note that here $\mathrm{E}[\cdot] = \int\int \cdot\ p(\boldsymbol{x}_-, \boldsymbol{u}|\boldsymbol{y}_-)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{u}$

$$= \int\int \cdot\ p(\boldsymbol{x}_-|\boldsymbol{y}_-)p(\boldsymbol{u})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{u} \qquad \text{(Markov Assumptions)}$$

2

Part II: Next, we get the mean $\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}|\boldsymbol{y}_-}$ of the marginal Gaussian $p(\boldsymbol{x}|\boldsymbol{y}_-) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-}, \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}|\boldsymbol{y}_-})$ from the joint Gaussian $p(\boldsymbol{x}, \boldsymbol{x}_-|\boldsymbol{y}_-) = p(\boldsymbol{w}|\boldsymbol{y}_-) \propto \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{w}}, \boldsymbol{\Sigma}_{\boldsymbol{w}\boldsymbol{w}})$. These parameters can be directly extracted from the mean $\boldsymbol{\mu}_{\boldsymbol{w}}$ and the covariance $\boldsymbol{\Sigma}_{\boldsymbol{w}\boldsymbol{w}}$ of the joint Gaussian (as derived in a previous class; see slides 5).

In particular, in this way we obtain

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}_-} = \boldsymbol{G}\boldsymbol{\mu}_{\boldsymbol{x}_-|\boldsymbol{y}_-} \qquad \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}|\boldsymbol{y}_-} = \boldsymbol{G}\boldsymbol{\Sigma}_{\boldsymbol{x}_-\boldsymbol{x}_-|\boldsymbol{y}_-}\boldsymbol{G}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{u}\boldsymbol{u}}$$

3

Derivation of Kalman Update Step

Recall update step of sequential Bayesian estimation

$$\underbrace{p(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{y}_-)}_{\text{Posterior pdf}} \propto p(\boldsymbol{y}|\boldsymbol{x}) \underbrace{p(\boldsymbol{x}|\boldsymbol{y}_-)}_{\substack{\text{Predicted}\\ \text{posterior pdf}}}$$

$$= p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{y}_-)p(\boldsymbol{x}|\boldsymbol{y}_-) \qquad \text{(Markov Assumptions)}$$

$$= p(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{y}_-)$$

Notation:

$\boldsymbol{x} = \boldsymbol{x}_n$

$\boldsymbol{y} = \boldsymbol{y}_n$

$\boldsymbol{v} = \boldsymbol{v}_n$

$\boldsymbol{x}_- = \boldsymbol{x}_{n-1}$

$\boldsymbol{y}_- = \boldsymbol{y}_{1:n-1}$

$\boldsymbol{H} = \boldsymbol{H}_n$

Note that $p(\boldsymbol{x}|\boldsymbol{y}_-) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}})$

Part I: First, we show that, given $\boldsymbol{y}_-$, $\boldsymbol{z} = [\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}}$ follows a joint Gaussian distribution, i.e.,

$$p(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{y}_-) = p(\boldsymbol{z}|\boldsymbol{y}_-) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{z}}, \boldsymbol{\Sigma}_{\boldsymbol{z}\boldsymbol{z}})$$

with mean and covariance given by

$$\boldsymbol{\mu}_{\boldsymbol{z}} = \begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{x}} \\ \boldsymbol{\mu}_{\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{x}} \\ \boldsymbol{H}\,\boldsymbol{\mu}_{\boldsymbol{x}} \end{bmatrix} \qquad \boldsymbol{\Sigma}_{\boldsymbol{z}\boldsymbol{z}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}} & \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{y}} \\ \boldsymbol{\Sigma}_{\boldsymbol{y}\boldsymbol{x}} & \boldsymbol{\Sigma}_{\boldsymbol{y}\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}} & \boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}}\boldsymbol{H}^{\mathrm{T}} \\ \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}} & \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{v}\boldsymbol{v}} \end{bmatrix}$$

4

Part I (cnt.): Note that $p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{y}_{-}) = p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{H}\boldsymbol{x}, \boldsymbol{\Sigma}_{\boldsymbol{v}})$ and $p(\boldsymbol{x}|\boldsymbol{y}_{-}) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{\Sigma}_{\boldsymbol{xx}})$ are both Gaussian distributed. Thus, we have that $p(\boldsymbol{z}|\boldsymbol{y}_{-}) = p(\boldsymbol{x},\boldsymbol{y}|\boldsymbol{y}_{-}) = p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{y}_{-})p(\boldsymbol{x}|\boldsymbol{y}_{-})$ follows a joint Gaussian distribution (as derived in a previous class; see slides 5). The missing parameters of the joint Gaussian distribution can be calculated as follows

$$\boldsymbol{\mu}_{\boldsymbol{y}} = \mathrm{E}(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{v}) = \boldsymbol{H}\boldsymbol{\mu}_{\boldsymbol{x}}$$

$$\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{yy}} &= \mathrm{E}\big[(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}})(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}})^{\mathrm{T}}\big] \\
&= \mathrm{E}\big[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\big] - \boldsymbol{\mu}_{\boldsymbol{y}}\boldsymbol{\mu}_{\boldsymbol{y}}^{\mathrm{T}} \\
&= \mathrm{E}\big[(\boldsymbol{H}\boldsymbol{x}+\boldsymbol{v})(\boldsymbol{H}\boldsymbol{x}+\boldsymbol{v})^{\mathrm{T}}\big] - \boldsymbol{H}\boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} \\
&= \boldsymbol{H}\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}]\boldsymbol{H}^{\mathrm{T}} + \mathrm{E}\big[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\big] - \boldsymbol{H}\boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} \\
&= \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{xx}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{vv}}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{yx}} &= \mathrm{E}\big[(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}})(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{x}})^{\mathrm{T}}\big] \\
&= \boldsymbol{H}\,\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}] - \mathrm{E}[\boldsymbol{y}]\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}} \\
&= \boldsymbol{H}(\boldsymbol{\Sigma}_{\boldsymbol{xx}} + \boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}}) - \boldsymbol{H}\boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}} \\
&= \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{xx}}
\end{aligned}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{yx}} = \boldsymbol{\Sigma}_{\boldsymbol{yx}}^{\mathrm{T}}$$

$$\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}] = \boldsymbol{\Sigma}_{\boldsymbol{xx}} + \boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{T}}$$

Note that here $\mathrm{E}[\cdot] = \int\int \cdot \; p(\boldsymbol{x},\boldsymbol{v}|\boldsymbol{y}_{-})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}$

$$= \int\int \cdot \; p(\boldsymbol{x}|\boldsymbol{y}_{-})p(\boldsymbol{v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v} \qquad \text{(Markov Assumptions)}$$

5

Part II: Next, we calculate the mean $\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}}$ of $p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{y}_{-}) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}}, \boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}})$ from $p(\boldsymbol{x},\boldsymbol{y}|\boldsymbol{y}_{-}) = p(\boldsymbol{z}|\boldsymbol{y}_{-}) \propto \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{z}}, \boldsymbol{\Sigma}_{\boldsymbol{zz}})$ by using the Schur complement (as derived in a previous class; see slides 5). In this way, we obtain

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu}_{\boldsymbol{x}} + \boldsymbol{\Sigma}_{\boldsymbol{xy}}\boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_{\boldsymbol{y}})$$

$$\boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}} = \boldsymbol{\Sigma}_{\boldsymbol{xx}} - \boldsymbol{\Sigma}_{\boldsymbol{xy}}\boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{yx}}$$

Equivalently, we can write

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu}_{\boldsymbol{x}} + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{\mu}_{\boldsymbol{y}})$$

$$\boldsymbol{\Sigma}_{\boldsymbol{xx}|\boldsymbol{y}} = \boldsymbol{\Sigma}_{\boldsymbol{xx}} - \boldsymbol{K}\boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{xx}}$$

where we have introduced the Kalman gain

$$\begin{aligned}
\boldsymbol{K} &= \boldsymbol{\Sigma}_{\boldsymbol{xy}}\boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1} \\
&= \boldsymbol{\Sigma}_{\boldsymbol{xx}}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{xx}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{v}})^{-1}
\end{aligned}$$

6