# 9.
# Recursive least squares

off-line version (with storing data in memory)

$$\widehat{a}_N = (X_N^T X_N)^{-1} X_N^T Y_N$$

on-line version (recursive)

$$a_N = f(a_{N-1}, x_N, y_N) = a_{N-1} + \delta(a_{N-1}, x_N, y_N)$$

Problem: matrix to be inverted

$$X_N^T X_N = \sum_{k=1}^{N} x_k x_k^T = \sum_{k=1}^{N-1} x_k x_k^T + x_N x_N^T = X_{N-1}^T X_{N-1} + x_N x_N^T$$

Let

$$P_N = (X_N^T X_N)^{-1}$$

$$cov(a_N) = P_N \sigma_z^2$$

We get

$$\begin{aligned} a_N &= P_N X_N^T Y_N \\ a_{N-1} &= P_{N-1} X_{N-1}^T Y_{N-1} \end{aligned}$$

$$P_N = \left(P_{N-1}^{-1} + x_N x_N^T\right)^{-1}, \text{ where } P_{N-1}^{-1} = X_{N-1}^T X_{N-1}$$

**Lemat 1** *It holds that*

$$\left(A + uu^T\right)^{-1} = A^{-1} - \frac{1}{1 + u^T A^{-1} u} A^{-1} u u^T A^{-1}$$

For $A = P_{N-1}^{-1}$ and $u = x_N$ we have that

$$\begin{aligned} P_N &= P_{N-1} - \frac{1}{1 + x_N^T P_{N-1} x_N} P_{N-1} x_N x_N^T P_{N-1} = P_{N-1} - \varkappa_N P_{N-1} x_N x_N^T P_{N-1} \\ \text{where } \varkappa_N &= \frac{1}{1 + x_N^T P_{N-1} x_N} \end{aligned}$$

$$
\begin{aligned}
a_N &= \left(P_{N-1} - \varkappa_N P_{N-1} x_N x_N^T P_{N-1}\right)\left(X_{N-1}^T Y_{N-1} + x_N y_N\right) = \\
&= P_{N-1} X_{N-1}^T Y_{N-1} + P_{N-1} x_N y_N - [\varkappa_N P_{N-1} x_N](x_N^T P_{N-1} X_{N-1}^T Y_{N-1} + x_N^T P_{N-1} x_N y_N) = \\
&= a_{N-1} + [\varkappa_N P_{N-1} x_N]\{\frac{1}{\varkappa_N} y_N - x_N^T a_{N-1} - x_N^T P_{N-1} x_N y_N\}
\end{aligned}
$$

but since

$$\frac{1}{\varkappa_N} = 1 + x_N^T P_{N-1} x_N$$

we conclude that

$$\{...\} = y_N + x_N^T P_{N-1} x_N y_N - x_N^T a_{N-1} - x_N^T P_{N-1} x_N y_N = y_N - x_N^T a_{N-1}$$

and hence

$$
\begin{aligned}
a_N &= a_{N-1} + \varkappa_N P_{N-1} x_N (y_N - x_N^T a_{N-1}) \\
P_N &= P_{N-1} - \varkappa_N P_{N-1} x_N x_N^T P_{N-1} \quad / \cdot x_N
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
P_N x_N &= P_{N-1} x_N - \varkappa_N P_{N-1} x_N x_N^T P_{N-1} x_N = \varkappa_N P_{N-1} x_N \{\frac{1}{\varkappa_N} - x_N^T P_{N-1} x_N\} \\
\{...\} &= 1 \\
P_N x_N &= \varkappa_N P_{N-1} x_N
\end{aligned}
$$

finally, (1) can be shown in the form

$$
\begin{aligned}
a_N &= a_{N-1} + P_N x_N (y_N - x_N^T a_{N-1}) \\
\text{where } P_N &= P_{N-1} - \frac{P_{N-1} x_N x_N^T P_{N-1}}{1 + x_N^T P_{N-1} x_N}
\end{aligned}
$$

initial conditions

$$a_0 = 0,\ P_0 = diag[10^3 \div 10^5]$$

**advantages: algorithm works without matrix inversion computation, measurement need not to be stored in memory**

weighted least squares (with exponential forgetting)

$$
\begin{aligned}
a_N &= a_{N-1} + P_N x_N (y_N - x_N^T a_{N-1}) \\
P_N &= \frac{1}{\lambda}\left(P_{N-1} - \frac{P_{N-1} x_N x_N^T P_{N-1}}{\lambda + x_N^T P_{N-1} x_N}\right)
\end{aligned}
$$

for dynamic system

$$
\begin{aligned}
\theta_N &= \theta_{N-1} + P_N \varphi_N (y_N - \varphi_N^T \theta_{N-1}) \\
P_N &= \frac{1}{\lambda}\left(P_{N-1} - \frac{P_{N-1} \varphi_N \varphi_N^T P_{N-1}}{\lambda + \varphi_N^T P_{N-1} \varphi_N}\right)
\end{aligned}
$$

instrumental variables

$$
\begin{aligned}
\theta_N &= \theta_{N-1} + P_N \psi_N (y_N - \varphi_N^T \theta_{N-1}) \\
P_N &= \frac{1}{\lambda}\left(P_{N-1} - \frac{P_{N-1} \psi_N \varphi_N^T P_{N-1}}{\lambda + \varphi_N^T P_{N-1} \psi_N}\right)
\end{aligned}
$$