LINEARIZATION

viates this problem by *linearizing* g and h via Taylor expansion—completely analogously to Equations (10.14) and (10.18) in the derivation of the EKF. In particular, we have:

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + G_t(x_{t-1} - \mu_{t-1}) \tag{11.16}$$

$$h(y_t, c_t^i) \approx h(\mu_t, c_t^i) + H_t^i\ (y_t - \mu_t) \tag{11.17}$$

Here μ_t is the current estimate of the state vector y_t, and $H_t^i = h_t^i\ F_{x,j}$ as defined already in Equation (10.19).

This linear approximation turns the log-likelihood (11.15) into a function that is quadratic in $y_{0:t}$. In particular, we obtain

$$\begin{aligned}\log p(y_{0:t} \mid z_{1:t}, u_{1:t}, c_{1:t}) &= \text{const.} - \tfrac{1}{2} \\ &\Big\{ x_0^T\ \Omega_0\ x_0 + \sum_t [x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})]^T \\ &\qquad R_t^{-1}\ [x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})] \\ &+ \sum_i [z_t^i - h(\mu_t, c_t^i) - H_t^i(y_t - \mu_t)]^T\ Q_t^{-1}\ [z_t^i - h(\mu_t, c_t^i) - H_t^i(y_t - \mu_t)] \Big\}\end{aligned} \tag{11.18}$$

This function is indeed a quadratic in $y_{0:t}$, and it shall prove convenient to reorder its terms, omitting several constant terms.

$$\begin{aligned}\log p(y_{0:t} \mid z_{1:t}, u_{1:t}, c_{1:t}) &= \text{const.} \\ &-\tfrac{1}{2} \underbrace{x_0^T\ \Omega_0\ x_0}_{\text{quadratic in } x_0} - \tfrac{1}{2} \sum_t \underbrace{x_{t-1:t}^T \begin{pmatrix} -G_t^T \\ 1 \end{pmatrix} R_t^{-1} \begin{pmatrix} -G_t & 1 \end{pmatrix} x_{t-1:t}}_{\text{quadratic in } x_{t-1:t}} \\ &+ \underbrace{x_{t-1:t}^T \begin{pmatrix} -G_t^T \\ 1 \end{pmatrix} R_t^{-1}\ [g(u_t, \mu_{t-1}) - G_t\ \mu_{t-1}]}_{\text{linear in } x_{t-1:t}} \\ &- \tfrac{1}{2} \sum_i \underbrace{y_t^T\ H_t^{iT}\ Q_t^{-1}\ H_t^i\ y_t}_{\text{quadratic in } y_t} + \underbrace{y_t^T\ H_t^{iT}\ Q_t^{-1}\ [z_t^i - h(\mu_t, c_t^i) + H_t^i \mu_t]}_{\text{linear in } y_t}\end{aligned} \tag{11.19}$$

Here $x_{t-1:t}$ denotes the state vector concatenating x_{t-1} and x_t; hence we can write $(x_t - G_t\ x_{t-1})^T = x_{t-1:t}^T\ (-G_t\ \ 1)^T = x_{t-1:t}^T \begin{pmatrix} -G_t^T \\ 1 \end{pmatrix}$.

If we collect all quadratic terms into the matrix Ω, and all linear terms into a vector ξ, we see that expression (11.19) is of the form

$$\log p(y_{0:t} \mid z_{1:t}, u_{1:t}, c_{1:t}) = \text{const.} - \tfrac{1}{2}\ y_{0:t}^T\ \Omega\ y_{0:t} + y_{0:t}^T\ \xi \tag{11.20}$$