Prove Kalman filter.

Let $$x_k = A_{k-1} x_{k-1} + q_{k-1}$$
$$y_k = H_k x_k + r_k.$$

$\rightarrow$ $q_{k-1} \sim N(0, Q_{k-1})$ white process noise.
$r_k \sim N(0, R_k)$ white measurement noise
A_{k-1} is the transition matrix.
H_k is measurement model matrix.

In probabilistic terms, model is
$$p(x_k | x_{k-1}) = N(x_k | A_{k-1} x_{k-1}, Q_{k-1})$$
$$p(y_k | x_k) = N(y_k | H_k x_k, R_k).$$

We know the Gaussian probability density is given by.

$$N(x | m, P) = \frac{1}{(2\pi)^{\frac{m}{2}} \; |P|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-m)^T P^{-1} (x-m)\right)$$

Let x & y have gaussian densities.

$$p(x) = N(x | m, P) \qquad p(y|x) = N(y | Hx, R)$$

Then joint & marginal distribution are

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left(\begin{pmatrix} m \\ Hm \end{pmatrix}, \begin{pmatrix} P & PH^T \\ HP & HPH^T + R \end{pmatrix}\right)$$

$$y \sim N(Hm, HPH^T + R)$$

We know that if random variables x & y have joint gaussian probability density.

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} A & C \\ C^{T} & B \end{pmatrix}\right),$$

Then marginal & conditional density for x & y are.

$$x \sim N(a, A)$$

$$y \sim N(b, B)$$

$$x|y \sim N(a + CB^{-1}(y-b), A - CB^{-1}C^{T})$$

$$y|x \sim N(b + C^{T}A^{-1}(x-a), B - C^{T}A^{-1}C).$$

Now since we assume a gaussian prior the posterior will also be gaussian, we get prediction as follows.

$$p(x_{k-1} | y_{1:k-1}) = N(x_{k-1} | m_{k-1}, P_{k-1})$$

$$p(x_k | y_{1:k-1}) = N(x_k | A_{k-1} m_{k-1}, A_{k-1} P_{k-1} A_{k-1}^{T} + Q_{k-1})$$

$$\boxed{f(y_k | x_{1:k}) = P(x_k | y_{1:k-1}) = N(x_k | m_{k|k}, P_{k|k})}$$

The joint distribution of y_k & x_k is

$$p(x_k, y_k \mid y_{1:k-1}) = p(y_k \mid x_k)\, p(x_k \mid y_{1:k-1})$$

$$= N\left(\begin{bmatrix} x_k \\ y_k \end{bmatrix} \middle| \begin{matrix} m_k \\ H_k m_k \end{matrix}, \begin{matrix} P_k & P_k H_k^T \\ H_k P_k & H_k P_k H_k^T + R_k \end{matrix} \right).$$

Thus the conditional distribution of x_k given y_k is given as.

$$p(x_k \mid y_k, y_{1:k-1}) = p(x_k \mid y_{1:k})$$

$$\boxed{f(y_T \mid x_{1:T}) = p(x_k \mid y_k, y_{1:k-1}) = N(x_k \mid m_k, P_k)}$$

where
$$P_k = P_{k-1} - K_k S_k K_k^T.$$
$$m_k = m_{k-1} + K_k (y_k - H_k m_{k-1})$$
$$K_k = P_{k-1} H_k^T S_k^{-1}$$
$$S_k = H_k P_{k-1} H_k^T + R_k.$$