(a) Ignoring the time index. $\tilde{y} = Cx + v$.

$$\hat{x} = \arg\min_x \|\tilde{y} - Cx\|_2^2$$
$$= \arg\min_x (\tilde{y} - Cx)^T(\tilde{y} - Cx)$$
$$= \arg\min_x \tilde{y}^T\tilde{y} + x^TC^TCx - x^TC^T\tilde{y} - \tilde{y}^TCx := g(x)$$

$$\frac{\partial g(x)}{\partial x} = 2C^TCx - 2C^T\tilde{y} = 0, \quad \frac{\partial g(x)}{\partial x^2} = 2C^TC \succ 0.$$

$\Rightarrow \hat{x} = (C^TC)^{-1}C^T\tilde{y}$. and this is minimum as Hessian is positive definite.

$$E(\hat{x}) = E[(C^TC)^{-1}C^T\tilde{y}] = (C^TC)^{-1}C^T E[\tilde{y}]$$
$$= (C^TC)^{-1}C^T E(Cx + v)$$
$$= (C^TC)^{-1}C^TC \cdot E(x) \qquad \text{as } E(v) = 0$$
$$= E(x)$$
$$= x. \qquad \square$$

(b) $$P := \mathrm{Cov}(x - \hat{x}) = E[(x-\hat{x}) - E(x-\hat{x}))(x-\hat{x}) - E(x-\hat{x}))^T]$$
$$= E[(x-\hat{x})\cdot(x-\hat{x})^T]$$
$$= E(xx^T + \hat{x}\hat{x}^T - x\hat{x}^T - \hat{x}x^T)$$
$$= xx^T + E[(C^TC)^{-1}C^T\tilde{y}\tilde{y}^TC(C^TC)^{-1}] - x\cdot E[(C^TC)^{-1}C^T\tilde{y}]^T - E[(C^TC)^{-1}C^T\tilde{y}]x^T$$
$$= xx^T + (C^TC)^{-1}C^T E(\tilde{y}\tilde{y}^T)\cdot C(C^TC)^{-1} - x[(C^TC)^{-1}C^T E[Cx+v]]^T - \cancel{E[Cx+v]}$$
$$- (C^TC)^{-1}C^T E[Cx+v]\cdot x^T$$

we have $$E[\tilde{y}\tilde{y}^T] = E[(Cx+v)(Cx+v)^T]$$
$$= E[Cx\cdot x^TC^T + Cxv^T + vx^TC^T + vv^T]$$
$$= C\cdot E[xx^T]\cdot C^T + R$$
$$= Cxx^TC^T + R.$$

therefore $$P = xx^T + (C^TC)^{-1}C^T[Cxx^TC^T + R]C(C^TC)^{-1} - xx^T - xx^T$$
$$= (C^TC)^{-1}C^TRC(C^TC)^{-1} \qquad \square$$

(c) $\hat{x}_2 = \underset{x}{\arg\min} \{ (x-\hat{x}_1)^T P_1^{-1} (x-\hat{x}_1) + (\tilde{y}-Cx)^T (\tilde{y}-Cx) \}.$

$= \underset{x}{\arg\min} \{ x^T P_1^{-1} x - x^T P_1^{-1} \hat{x}_1 - \hat{x}_1^T P_1^{-1} x + \hat{x}_1^T P_1^{-1} \hat{x}_1 + (\tilde{y}-Cx)^T (\tilde{y}-Cx) \} := g(x)$

$\Rightarrow \nabla_x g(x) = 2P_1^{-1} x - 2P_1^{-1}\hat{x}_1 - 2C^T(\tilde{y}-Cx) = 0$

$\Rightarrow (P_1^{-1} + C^T C) x = P_1^{-1}\hat{x}_1 + C^T \tilde{y}.$

$\Rightarrow \hat{x}_2 = (P_1^{-1}+C^TC)^{-1} P_1^{-1} \hat{x}_1 + (P_1^{-1}+C^TC)^{-1} C^T \tilde{y}.$

$\nabla^2_x g(x) = 2(P_1^{-1}+C^TC) \succ 0.$

therefore. $\hat{x}_2 = (P_1^{-1}+C^TC)^{-1} P_1^{-1} \hat{x}_1 + (P_1^{-1}+C^TC)^{-1} C^T \tilde{y}.$

$\Rightarrow \Sigma = (P_1^{-1}+C^TC)^{-1} P_1^{-1}, \quad \beta = (P_1^{-1}+C^TC)^{-1} C^T \tilde{y}$

(d) Based on the prior distribution $N(\hat{x}_1, P_1)$.

the distribution density is $p \propto e^{-\frac{1}{2}(x-\hat{x}_1)^T P_1^{-1}(x-\hat{x}_1)}.$

we hope the ~~posterior is~~ new state $\hat{x}_2$ as close to its prior as possible, therefore $\max.\ e^{-\frac{1}{2}(x-\hat{x}_1)^T P_1^{-1}(x-\hat{x}_1)}$

$\Leftrightarrow \min (x-\hat{x}_1)^T P_1^{-1}(x-\hat{x}_1).$

Hence we add the first term.

(e) If $\hat{x}^+ = L\hat{x}^- + K\tilde{y}$. as $\hat{x}^+$ is unbiased.

then $\mathbb{E}(\hat{x}^+) = x$

$\Rightarrow \mathbb{E}[L\hat{x}^- + K\tilde{y}]$

$= L \cdot \mathbb{E}(\hat{x}^-) + K\,\mathbb{E}[Cx+v]$

$= L \cdot x + K \cdot C \cdot x$

$= (L+KC) x$

$= x$

$\Rightarrow L + KC = I$

$\Rightarrow \hat{x}^+ = (I-KC)\hat{x}^- + K\tilde{y} = \hat{x}^- + K(\tilde{y} - C\hat{x}^-) \quad \square$

(f) As $\hat{x}^+$ is unbiased. $\mathbb{E}(x-\hat{x}^+)=0$

$$\Rightarrow \quad P^+ = \operatorname{cov}(x-\hat{x}^+) = \mathbb{E}[(x-\hat{x}^+)(x-\hat{x}^+)^T].$$

By property of traces. $\operatorname{Tr}(ab^T) = a^T b$.

we have.

$\cancel{\operatorname{Tr}[P^+] = \operatorname{Tr}}$

$$\operatorname{Tr}[(x-\hat{x}^+)(x-\hat{x}^+)^T] = (x-\hat{x}^+)^T(x-\hat{x}^+) = \|x-\hat{x}^+\|_2^2.$$

Take expectation

$$\mathbb{E}\{\operatorname{Tr}[(x-\hat{x}^+)(x-\hat{x}^+)^T]\} = \mathbb{E}[\|x-\hat{x}^+\|_2^2].$$

Because Tr is a linear operator, we can push expectation insider.

$$\Rightarrow \quad \operatorname{Tr}\left[\mathbb{E}[(x-\hat{x}^+)(x-\hat{x}^+)^T]\right] = \mathbb{E}[\|x-\hat{x}^+\|_2^2]$$

$$\Leftrightarrow \quad \operatorname{Tr}[P^+] = \mathbb{E}[\|x-\hat{x}^+\|_2^2]. \qquad \square$$

(g)

$$\begin{aligned} x-\hat{x}^+ &= x-(\hat{x}^- + K(y - C\hat{x}^-)) \\ &= x-\hat{x} - K\{C(x-\hat{x}^-)+v\} \\ &= (I-KC)(x-\hat{x}^-) - Kv. \end{aligned}$$

$$\cancel{\|x-\hat{x}^+\|_2^2 = (x-\hat{x}^-)^T(I-KC)^T(I-KC)(x-\hat{x}^-) + v^TK^TKv - (x-\hat{x}^-)^T(I-KC)^TKv - v^TK^T(I-KC)(x-\hat{x}^-)}$$

$$\cancel{\operatorname{Tr}[P^+] = \mathbb{E}[\|x-\hat{x}^+\|_2^2]}$$

$$\begin{aligned} P^+ &= \mathbb{E}[(x-\hat{x}^+)(x-\hat{x}^+)^T] \\ &= \mathbb{E}\{[(I-KC)(x-\hat{x}^-)-Kv][(I-KC)(x-\hat{x}^-)-Kv]^T\}. \\ &= (I-KC)\,\mathbb{E}[(x-\hat{x}^-)(x-\hat{x}^-)^T](I-KC)^T + KRK^T. \\ &= (I-KC)P^-(I-KC)^T + KRK^T. \\ &= P^- - KCP^- - P^-C^TK^T \cancel{+ C^TK^TPK} + KCP^-C^TK^T + KRK^T \end{aligned}$$

$$\operatorname{Tr}(P^+) = \operatorname{Tr}(P^-) - \operatorname{Tr}(KCP^-) - \operatorname{Tr}(P^-C^TK^T) + \operatorname{Tr}(KCP^-C^TK^T) + \operatorname{Tr}(KRK^T)$$

By Matrix cookbook. (wikipedia.)

$\frac{\partial}{\partial X} Tr(XA) = A^T$, $\frac{\partial}{\partial X} Tr(XBX^T) = XB^T + XB$, $\frac{\partial}{\partial X} Tr(AX^T) = A$.

We then have

$$\frac{\partial}{\partial K} Tr(P^+) = -P^{-T}C^T - P^-C^T + K(CP^-C^T+R)^T + K\cdot(CP^-C^T+R) = 0$$

Because R, P is covariance matrix. $P^- = P^{-T}$, $R = R^T$

$$\Rightarrow \quad K\cdot(CP^-C^T+R)^T = P^-C^T$$

$$K = P^-C^T(CP^-C^T+R)^{-1} \qquad \square$$

(h).
$$P^+ = (I-KC)P^-(I-KC)^T + KRK^T$$
$$= (I-KC)P^- - (I-KC)P^-C^T\underline{K^T} + KR\underline{K^T}$$
$$= (I-KC)P^- + [(I-KC)P^-C^T + P^-C^T(CP^-C^T+R)^{-1}R]K^T$$
$$= (I-KC)P^- + [-\{I-P^-C^T(CP^-C^T+R)^{-1}C\}P^-C^T + P^-C^T(CP^-C^T+R)^{-1}R]K^T$$
$$= (I-KC)P^- + [P^-C^T(CP^-C^T+R)^{-1}[CP^-C^T+R] - P^-C^T]K^T$$
$$= (I-KC)P^- \qquad \square$$

(i) As $\hat{x}^-_{k+1} = A\hat{x}^+_k$, if ~~$\hat{x}^+_k$ is unbiased, then so is $\hat{x}^-_{k+1}$~~.

$$P^-_{k+1} = E[(\hat{x}^-_{k+1} - x_{k+1})(\hat{x}^-_{k+1} - x_{k+1})^T]$$
$$\cancel{= E[\hat{x}^-_{k+1}\hat{x}^{-T}_{k+1}]}$$
$$= E[[A\hat{x}^+_k - (Ax_k+w_k)][A\hat{x}^+_k - (Ax_k+w_k)]^T]$$
$$= A\,E[(\hat{x}^+_k - x_k)(\hat{x}^+_k - x_k)^T]A^T + E[w_k w_k^T] \quad \text{as } \hat{x}^+_k \text{ is unbiased.}$$
$$= AP^+_kA^T + Q. \qquad \square$$

(j)

1° update $\hat{x}^-_{k+1} = A\hat{x}^+_k$, $P^-_{k+1} = AP^+_kA^T + Q$

2° Compute $K_{k+1} = P^-_{k+1}C^T(CP^-_{k+1}C^T+R)^{-1}$.

3° update. $\hat{x}^+_{k+1} = \hat{x}^-_{k+1} + K(\tilde{y} - C\hat{x}^-_{k+1})$. $P^+_{k+1} = (I-K_{k+1}C)P^-_{k+1}$.