Madhushree Sannigrahi

Student no. – 128570

Sensor fusion – II

Detailed Feedback

Task 1: *Transformation of Gaussian random variables*

Let $x \in \mathbb{R}^n$ be $\mathcal{N}(\mu, \Sigma)$. Find the distribution and see if you recognize it:

Hint: they are all given in the book.

(a) $z = \Sigma^{-\frac{1}{2}}(x - \mu)$, where $\Sigma^{\frac{1}{2}}(\Sigma^{\frac{1}{2}})^T = \Sigma$

> *Hint:* If you are using theorem 2.4.1, you might need $\det(A^{\frac{1}{2}}) = \det(A)^{\frac{1}{2}}$, $(A^{-1})^T = (A^T)^{-1}$, and $\det(A^T) = \det(A)$ whenever A has full rank.

$$x \sim \mathcal{N}(\mu, \Sigma)$$

$$z = \Sigma^{-1/2}(x - \mu) \quad \text{where } \Sigma = \Sigma^{1/2}(\Sigma^{1/2})^T$$

$$g(x) \propto \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x - \mu)\right)$$

$$h(z) \propto \exp\left(-\frac{1}{2}(\Sigma^{1/2} z)^T \Sigma^{-1} (\Sigma^{1/2} z)\right)$$

$$= \exp\left(-1/2\, (\Sigma^{1/2} z)^T \cancel{(\Sigma^{1/2})^T} \cancel{((\Sigma^{1/2})^T)^{-1}} \Sigma^{-1} \cancel{(\Sigma^{1/2})^{-1} (\Sigma^{1/2})} (\Sigma^{1/2} z)\right)$$

$$= \exp\left(-1/2\; z^T \underline{(\Sigma^{1/2})^T \Sigma^{-1} \Sigma^{1/2}} z\right)$$ [symmetric positive definite]

$$= \exp\left(-1/2\; z^T I z\right)$$

$z \sim \mathcal{N}(0, I)$ [whitening in Gaussian distribution]

$$z = \Sigma^{-1/2}(x - \mu)$$
$$= \Sigma^{-1/2} x - \Sigma^{-1/2} \mu$$
$$\therefore x = \frac{\Sigma^{-1/2}\mu + z}{\Sigma^{-1/2}} = \Sigma^{1/2} z + \mu$$

(b) Use transformation of random variables to find $y_i = z_i^2$, where z_i is the i'th variable in the vector z.

$$y_i = z_i^2 \qquad \therefore\ z_i = \underbrace{\sqrt{y_i}}_{f_1^{-1}} \text{ or } \underbrace{-\sqrt{y}}_{f_2^{-1}}$$

Finding the Jacobians,

$$f_1^{-1}(y_i) = \frac{1}{2\sqrt{y_i}} \text{ and } f_2^{-1}(y_i) = -\frac{1}{2\sqrt{y_i}}$$

Stitching together, we get:

$$m(y_i) = \frac{1}{2\pi}\left[\frac{1}{2\sqrt{y_i}} e^{-\frac{1}{2}(\sqrt{y_i})^2} + \frac{1}{2\sqrt{y_i}} e^{-\frac{1}{2}(-\sqrt{y_i})^2}\right] = \frac{1}{\sqrt{2\pi y_i}} e^{-y/2}$$ [χ^2 distribution with 1 DOF]

(c) $y = (x-\mu)^T \Sigma^{-1} (x-\mu) = z^T z = \sum z_i^2 = \sum y_i$.

$$M(s) = \int_0^{\infty} e^{sy}\, m(y_i)\, dy_i \quad \left[\text{Moment generating function}\right]$$

$$= \int_0^{\infty} e^{sy} \frac{1}{\sqrt{2\pi y_i}}\, e^{-y/2}\, dy_i$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{e^{(s-1/2)y_i}}{\sqrt{y_i}}\, dy_i$$

taking $s - \frac{1}{2} = -a$ and,

$$y_i = z_i^2$$
$$dy_i = 2 z_i\, dz_i$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-a z_i^2}\, dz_i$$

$$\left[\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}\right]$$

$$= \frac{\cancel{2}}{\sqrt{2}\,\cancel{\sqrt{\pi}}} \cdot \frac{1}{\cancel{2}} \frac{\cancel{\sqrt{\pi}}}{\sqrt{-a}}$$

$$= \frac{1}{\sqrt{2\left(\frac{1}{2}-s\right)}} = \frac{\sqrt{1/2}}{\sqrt{1/2 - s}}$$

$$M_y(s) = \left[M(s)\right]^n = \left[\frac{\sqrt{1/2}}{\sqrt{1/2 - s}}\right]^n = \left[\frac{1/2}{1/2 - s}\right]^{n/2}$$

Performing inverse Laplace transform, we get:

$$y^{n/2}\, e^{-\frac{1}{2}y}$$

$$\therefore\ y \sim \frac{y^{n/2-1}\, e^{-y/2}}{2^{n/2}\, \Gamma(n/2)} \quad \left[\text{Using Example 2.8}\right]$$

Gamma distribution with scale parameter : 2
shape parameter : $n/2$

Task 2: *Sensor fusion*

In this task we want to find out if a boat is above the line $x_2 = x_1 + 5$. In order to do this we will fuse measurements from two sensors with our prior belief: A drone-mounted camera, and a maritime surveillance radar. You have some prior knowledge of the state of the boat. You get 1 measurement from each sensor that are processed so that you know them to be (approximately) Gaussian conditioned on the position.

To be more specific, let us denote the state by x and our prior Gaussian by $\mathcal{N}(x; \bar{x}, P)$. The measurement from the camera is given by $z^c = H^c x + v^c$ and the measurement from the radar by $z^r = H^r x + v^r$, where v^c, v^r denotes the measurement noise and is distributed according to $\mathcal{N}(0, R^c)$ and $\mathcal{N}(0, R^r)$, respectively.

Only insert the numbers when asked to. The needed values are given by

$$\bar{x} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T, \qquad P = 25I_2, \qquad H^c = H^r = I_2,$$

$$R^c = \begin{bmatrix} 79 & 36 \\ 36 & 36 \end{bmatrix}, \qquad R^r = \begin{bmatrix} 28 & 4 \\ 4 & 22 \end{bmatrix}, \qquad z^c = \begin{bmatrix} 2 & 14 \end{bmatrix}^T, \qquad z^r = \begin{bmatrix} -4 & 6 \end{bmatrix}^T$$

(a) What is $p(z^c|x)$?

(b) Show that the joint $p(x, z^c)$ can be written as a Gaussian distribution.

> *Hint:* Use conditional probability and the proof of theorem 3.3.1.

$$p(x) \sim \mathcal{N}(x; \bar{x}, P)$$

$$v^c \sim \mathcal{N}(0, R^c)$$

$$z^c = H^c x + v^c \quad \therefore \; H^c x \sim \mathcal{N}(x; H^c\bar{x}, H^c P H^{cT})$$

[Linearity]

$$\therefore \quad z^c \sim \mathcal{N}(z^c; H^c\bar{x}, H^c P H^{cT} + R^c)$$

[sum of variance]

$$\because \; Cov(a, b) = E\left[(a - E(a))(b - E(b))^T\right]$$

$$P_{(z^c\, x)} = Cov(H_c x + v_c, x)$$

Since v^c is independent of x & $\mu_{v^c} = 0$,

$$P_{(z^c, x)} = Cov(H^c x, x) = H^c Cov(x)$$
$$= H^c P$$

b)

$$p(z^c, x) = \mathcal{N}\left(\begin{bmatrix} z^c \\ x \end{bmatrix}; \begin{bmatrix} H^c\bar{x} \\ \bar{x} \end{bmatrix}, \begin{bmatrix} H^c P H^{cT} + R^c & H^c P \\ P H^{cT} & P \end{bmatrix}\right)$$

$$\mu_{(z^c|x)} = H^c\bar{x} + \cancel{H^c P}\, \cancel{P^{-1}}(x-\bar{x}) \qquad P_{(z_c|x)} = \cancel{H^c P H^{cT}} + R^c - \cancel{H^c P P^{-1} P H^{cT}}$$
$$= \cancel{H^c\bar{x}} + H^c x - \cancel{H^c\bar{x}} \qquad = R^c$$
$$= H^c x$$

a) $\therefore\ p(z^c|x) \sim \mathcal{N}(z^c;\ H^c x,\ R^c)$

(c) Find the marginal $p(z^c)$ and the conditional $p(x|z^c)$, using the above and either theorems from the book or calculations.

$$p(x|z^c) = \frac{p(z^c|x)\ p(x)}{p(z^c)} \qquad \text{[Using Bayes Th.]}$$

$$= \frac{\mathcal{N}(z^c;\ H^c x,\ R_c) \cdot \mathcal{N}(x;\ \bar{x}, P)}{\mathcal{N}(z^c;\ H^c\bar{x},\ H^c P H^{cT} + R^c)}$$

Following product identity rule,

$$p(x|z^c) = (z^c - H^c x)^T R^{c^{-1}} (z^c - H^c x) + (x-\bar{x})^T P^{-1}(x-\bar{x}) - (z^c - H^c\bar{x})^T (H^c P H^{cT} + R^c)^{-1}(z^c - H^c\bar{x})$$

$$= \begin{bmatrix} z^c - H^c x \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}^T \begin{bmatrix} R^{c^{-1}} & 0 & 0 \\ 0 & P^{-1} & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c)^{-1} \end{bmatrix} \begin{bmatrix} z^c - H^c x \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}$$

$$= \begin{bmatrix} z^c - H^c\bar{x} \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}^T \begin{bmatrix} I & 0 & 0 \\ -H^c & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} R^{c^{-1}} & 0 & 0 \\ 0 & P^{-1} & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c)^{-1} \end{bmatrix} \begin{bmatrix} I & -H^c & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} z^c - H^c x \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}$$

$$\begin{bmatrix} I & -H^c & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} z^c - H^c\bar{x} \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix} = \begin{bmatrix} z^c - H x + H^c\bar{x} - H^c\bar{x} \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix} = \begin{bmatrix} z^c - H^c x \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}$$

$$= \begin{bmatrix} z^c - H^c\bar{x} \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}^T \left(\begin{bmatrix} I & H^c & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} R^c & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c) \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ H^c & I & 0 \\ 0 & 0 & I \end{bmatrix} \right)^{-1} \begin{bmatrix} z^c - H^c\bar{x} \\ x - \bar{x} \\ z^c - H^c\bar{x} \end{bmatrix}$$

$$= \left(\begin{bmatrix} R & H^c P & 0 \\ 0 & P & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c) \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ H^c & I & 0 \\ 0 & 0 & I \end{bmatrix} \right)^{-1}$$

$$= \begin{bmatrix} z^c - H^c \bar{x} \\ x - \bar{x} \\ z^c - H^c \bar{x} \end{bmatrix}^T \left(\begin{bmatrix} R^c + H^c P H^c & P H^c & 0 \\ P H^c & P & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c) \end{bmatrix} \right)^{-1} \begin{bmatrix} z^c - H^c \bar{x} \\ x - \bar{x} \\ z^c - H^c \bar{x} \end{bmatrix}$$

$$p(z^c) \sim \mathcal{N}(z^c; H^c \bar{x}, H^c P H^{cT} + R^c) \text{ [from 2(a)]}$$

$$p(x \mid z^c) \sim \mathcal{N}\left(\begin{bmatrix} z^c \\ x \\ z^c \end{bmatrix}; \begin{bmatrix} H^c \bar{x} \\ \bar{x} \\ H^c \bar{x} \end{bmatrix}, \begin{bmatrix} R^c + H^c P H^c & P H^c & 0 \\ P H^c & P & 0 \\ 0 & 0 & -(H^c P H^{cT} + R^c) \end{bmatrix} \right)$$

(d) Given what you found above, what is the marginal $p(z^r)$ and the conditional $p(x|z^r)$?

$$p(z^r) \sim \mathcal{N}(z^r; H^r \bar{x}, H^r P H^{rT} + R^r)$$

$$p(x \mid z^r) \sim \mathcal{N}\left(\begin{bmatrix} z^r \\ x \\ z^r \end{bmatrix}; \begin{bmatrix} H \bar{x} \\ \bar{x} \\ H \bar{x} \end{bmatrix}, \begin{bmatrix} R^r + H^r P H^r & P H^r & 0 \\ P H^r & P & 0 \\ 0 & 0 & -(H^r P H^{rT} + R^r) \end{bmatrix} \right)$$

(e) What is the MMSE and MAP estimate of x given z^c? You do not need to do calculations to find the answer, but briefly state what you would do if you had to.

$$\hat{x}_{MMSE} = E\left[x \mid z^c\right] = \int x \, p(x \mid z^c) \, dx$$

we can substitute $p(x|z^c)$ with its distribution in the above equation and solve it using change of variable technique.

$$\hat{x}_{MAP} = \underset{x}{\arg\max}\; p(x \mid z^c) = \underset{x}{\arg\max}\; p(z^c \mid x) \, p(x)$$

we can solve it by multiplying the terms and then differentiating the log of the function w.r.t x. Taking the equivalent of this equation as 0 gives us the MAP of x given z^c.

(f) Finish the `sensor_model.LinearSensorModel2d.get_pred_meas` method that can be used to calculate marginal probabilities $p(z)$.

```
def get_pred_meas(self, state_est: MultiVarGauss2d) -> MultiVarGauss2d:
    pred_mean = self.H @ state_est.mean
    pred_cov = (self. H @ state_est.cov @ self.H.T) + self.R

    pred_meas = MultiVarGauss2d(pred_mean, pred_cov)

    return pred_meas
```

(g) Finish the `conditioning.get_cond_state` function that can be used to calculate conditional probabilities $p(x|z)$.

```
def get_cond_state(state: MultiVarGauss2d,
                   sens_modl: LinearSensorModel2d,
                   meas: Measurement2d
                   ) -> MultiVarGauss2d:

    pred_meas =  np.linalg.inv(sens_modl.H.T @ np.linalg.inv(sens_modl.R) @ sens_modl.H + np.linalg.inv(state.cov))
    kalman_gain =  pred_meas @ sens_modl.H.T @ np.linalg.inv(sens_modl.R)
    innovation =  meas.value - sens_modl.H.T @ state.mean
    cond_mean =  state.mean + kalman_gain @ innovation
    p,q = np.shape(kalman_gain @ sens_modl.H)
    cond_cov =   (np.eye(p,q) - kalman_gain @ sens_modl.H) @ state.cov

    cond_state = MultiVarGauss2d(cond_mean, cond_cov)

    return cond_state
```

(h) Finish the `task2.get_conds` function that is used to calculate the conditional probabilities $p(x|z^c)$ and $p(x|z^r)$.

(i) Finish the `task2.get_double_conds` function that is used to calculate the conditional probabilities $p(x|z^c, z^r)$, i.e. the posterior of x conditioned on z^c then z^r, and $p(x|z^r, z^c)$, i.e. the posterior of x conditioned on z^r then z^c. Does it matter which order we condition?

```
def get_conds(state: MultiVarGauss2d,
              sens_model_c: LinearSensorModel2d, meas_c: Measurement2d,
              sens_model_r: LinearSensorModel2d, meas_r: Measurement2d
              ) -> Tuple[MultiVarGauss2d, MultiVarGauss2d]:

    cond_c = get_cond_state(state,sens_model_c, meas_c)
    cond_r = get_cond_state(state,sens_model_r, meas_r)

    return cond_c, cond_r


def get_double_conds(state: MultiVarGauss2d,
                     sens_model_c: LinearSensorModel2d, meas_c: Measurement2d,
                     sens_model_r: LinearSensorModel2d, meas_r: Measurement2d
                     ) -> Tuple[MultiVarGauss2d, MultiVarGauss2d]:

    cond_c, cond_r = get_conds(state, sens_model_c, meas_c, sens_model_r, meas_r)
    cond_cr = get_cond_state(cond_c,sens_model_r, meas_r )
    cond_rc = get_cond_state(cond_r,sens_model_c, meas_c )

    return cond_cr, cond_rc
```

```
cond_c=
|-1.892e+00 |      | 1.745e+01 | 4.457e+00|
| 6.854e+00 |      | 4.457e+00 | 1.212e+01|

 cond_r=
|-2.141e+00 |      | 1.313e+01 | 1.010e+00|
| 3.374e+00 |      | 1.010e+00 | 1.162e+01|

cond_cr=
|-2.718e+00 |      | 1.070e+01 | 2.326e+00|
| 6.487e+00 |      | 2.326e+00 | 7.768e+00|

 cond_rc=
|-2.718e+00 |      | 1.070e+01 | 2.326e+00|
| 6.487e+00 |      | 2.326e+00 | 7.768e+00|
```

Order doesn't matter!! It is the same thing!

(j) Finish the `gaussian.MultiVarGauss2d.get_transformed` method that is used to calculate the probability $p(Tx)$, where T is a linear transformation.

```
def get_transformed(self, lin_transform: np.ndarray) -> 'MultiVarGauss2d':
    transformed_mean = self.mean @ lin_transform
    transformed_cov = lin_transform @ self.cov @ lin_transform.T
    transformed = MultiVarGauss2d(transformed_mean, transformed_cov)

    return transformed
```

(k) You now want to know the probability that the boat is above the line, $x_2 = x_1 + 5$. Finish `task2.get_prob_over_line` using the appropriate linear transform and the CDF.

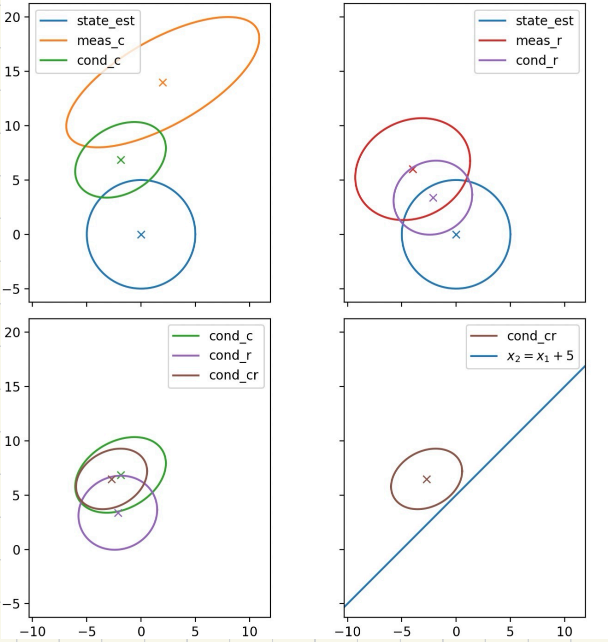
```
def get_prob_over_line(gauss: MultiVarGauss2d) -> float:

    # for x_2 = x_1 + 5

    lin_trans = np.array([-1, 1])
    transform  = gauss.get_transformed(lin_transform=lin_trans)
    prob = norm.cdf(5, transform.mean, np.sqrt(transform.cov))

    return prob
```

```
Prob that cond_cr is above x_2 = x_1 + 5 is 0.1289691345707723
```



Task 3: *Working with the canonical form*

In Section 3.3 the fundamental product identity was studied using a moment-based parametrization. Clearly, it must also be possible to establish an equivalent result using the canonical representation. In this exercise we shall therefore consider the product

$$\mathcal{N}^{-1}(\mathbf{x}\,;\,\mathbf{a},\mathbf{B})\mathcal{N}^{-1}(\mathbf{y}\,;\,\mathbf{C}\mathbf{x},\mathbf{D}). \tag{6}$$

(a) Show that (6) is identical to the Gaussian

$$\mathcal{N}^{-1}\left(\begin{bmatrix}\mathbf{x}\\\mathbf{y}\end{bmatrix};\begin{bmatrix}\mathbf{a}\\\mathbf{0}\end{bmatrix},\begin{bmatrix}\mathbf{B}+\mathbf{C}^{\mathsf{T}}\mathbf{D}^{-1}\mathbf{C} & -\mathbf{C}^{\mathsf{T}}\\-\mathbf{C} & \mathbf{D}\end{bmatrix}\right) \tag{7}$$

$$\mathcal{N}(x;\mu,P) = \exp\left(a + \eta^T x - \frac{1}{2}x^T \Lambda x\right)$$

where, $\Lambda = P^{-1}$, $\eta = \Lambda\mu$, $a = -\frac{1}{2}\left(n\ln(2\pi) - \ln|\Lambda| + \eta^T\Lambda^{-1}\eta\right)$

[canonical form]

$$\mathcal{N}_1^{-1}(x;\,a,\,B) = \exp\left(p + a^T x - \frac{1}{2}x^T Bx\right)$$

$$\mathcal{N}_2^{-1}(y;\,Cx,\,D) = \exp\left(q + C^T x^T y - \frac{1}{2}y^T Dy\right)$$

$$p(x,y) = \mathcal{N}_1^{-1}\mathcal{N}_2^{-1}$$

$$= \exp\left(p + q + a^T x + C^T x^T y - \frac{1}{2}x^T Bx - \frac{1}{2}y^T Dy\right)$$

$$= \exp\Big(\underbrace{-n\ln(2\pi) - \ln|B| + a^T B^{-1} a - \ln|D|}_{\gamma} - x^T C^T D^{-1} C x$$
$$+ a^T x + C^T x^T y - \frac{1}{2}x^T Bx - \frac{1}{2}y^T Dy\Big)$$

$$= \exp\left(\gamma + a^T x - \frac{1}{2}\left(x^T(B + C^T D^{-1} C)x - 2C^T x^T y + y^T D y\right)\right)$$

$$= \mathcal{N}^{-1}\left(\begin{bmatrix}x\\y\end{bmatrix};\begin{bmatrix}a\\0\end{bmatrix},\begin{bmatrix}B+C^T D^{-1}C & -C^T\\-C & D\end{bmatrix}\right)$$

(b) Show that the marginal distribution of $\mathbf{y}$, from the joint density (7), is

$$\mathcal{N}^{-1}\left(\mathbf{y}\,;\,\mathbf{C}^\top(\mathbf{B}+\mathbf{C}^\top\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{a}, \mathbf{D}-\mathbf{C}(\mathbf{B}+\mathbf{C}^\top\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{C}^\top\right).$$

$$P(y) = \mathcal{N}^{-1}(y;\, \eta_*, \Lambda_*)$$

$$\eta_* = 0 + C^T(B + C^T D^{-1} C)^{-1} a$$

[marginalization in canonical form]

$$\Lambda_* = D - C(B + C^T D^{-1} C)^{-1} C^T$$

(c) Show that the conditional distribution of $\mathbf{x}$ given $\mathbf{y}$ is

$$\mathcal{N}^{-1}\left(\mathbf{x}\,;\,\mathbf{a}+\mathbf{C}^\top\mathbf{y}, \mathbf{B}+\mathbf{C}^\top\mathbf{D}^{-1}\mathbf{C}\right).$$

$$P(x|y) = \mathcal{N}^{-1}(x;\, \eta_{x|y}, \Lambda_{x|y})$$

[conditioning in canonical form]

$$\eta_{x|y} = a + C^T y$$

$$\Lambda_{x|y} = B + C^T D^{-1} C$$

(d) Let us now return to the original formulation of the product identity in Theorem 3.3.1. Use the result from c) to show that

$$\widehat{\mathbf{P}}^{-1} = \mathbf{P}^{-1} + \mathbf{H}^\top\mathbf{R}^{-1}\mathbf{H}. \tag{10}$$

$$p(x|y) \sim \mathcal{N}^{-1}(x;\, a + C^T y,\, B + C^T D^{-1} C)$$

$$= \mathcal{N}^{-1}(x;\, (B + C^T D^{-1} C)^{-1}(a + C^T y);\, (B + C^T D^{-1} C)^{-1})$$

According to 3.16 in book,

$$\hat{P} = P_{x|y} = (B + C^T D^{-1} C)^{-1}$$

$$\hat{P}^{-1} = B + C^T D^{-1} C$$

Comparing it with 3.10,

$$\mathcal{N}^{-1}(x; a, B)\, \mathcal{N}^{-1}(y; Cx, D)$$

$$= \mathcal{N}(x; B^{-1}a, B^{-1})\, \mathcal{N}(y; D^{-1}Cx, D^{-1}) \quad \cdots\cdots (i)$$

Comparing (i) with $\mathcal{N}(x; \bar{x}, P)\mathcal{N}(z; Hx, R)$

we get the RHS of the given eqn.

$$P^{-1} + H^T R^{-1} H$$

$$= (B^{-1})^{-1} + (D^{-1}C)^T D (D^{-1}C)$$

$$= B + C^T D^{-1} C = \hat{P} : LHS$$