EE 150 - Applications of Convex Optimization in Signal Processing and Communications
Dr. Andre Tkacenko, JPL
Third Term 2011-2012

Homework Set #1 - Solutions

1. This is self-explanatory.

2. (a) Clearly the inequality holds if $a = 0$ or $b = 0$. Furthermore, in this case, it is easy to see that we have equality if and only if $a^p = b^q = 0$. When $a, b > 0$, we have the following:

$$\begin{aligned} ab &= e^{\log(ab)} = e^{[\log(a)+\log(b)]} = e^{\left[\frac{p}{p}\log(a)+\frac{q}{q}\log(b)\right]}, \\ &= e^{\left[\frac{1}{p}\log(a^p)+\frac{1}{q}\log(b^q)\right]}, \\ &\leq \frac{1}{p}e^{\log(a^p)} + \frac{1}{q}e^{\log(b^q)}, \qquad (1) \\ &= \frac{a^p}{p} + \frac{b^q}{q}. \end{aligned}$$

Here, (1) follows from the hint with $\theta = \frac{1}{p}$. Assuming $p < \infty$ (the special case $p = \infty$ can be argued via a continuity argument as $p \to \infty$), it follows that $0 < \theta < 1$, and so we have equality in (1) if and only if $\log(a^p) = \log(b^q)$, which is equivalent to saying $a^p = b^q$.

(b) When either $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$, the inequality clearly holds with equality if and only if $\mathbf{x} = \mathbf{y} = \mathbf{0}$, which is consistent with the condition for equality stipulated in the problem. For the case where $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, we have $||\mathbf{x}||_p > 0$ and $||\mathbf{y}||_q > 0$ and so we get the following for any k with $1 \leq k \leq n$:

$$\frac{|x_k y_k|}{||\mathbf{x}||_p\,||\mathbf{y}||_q} = \frac{|x_k|}{||\mathbf{x}||_p} \cdot \frac{|y_k|}{||\mathbf{y}||_q} \leq \frac{1}{p} \cdot \frac{|x_k|^p}{||\mathbf{x}||_p^p} + \frac{1}{q} \cdot \frac{|y_k|^q}{||\mathbf{y}||_q^q}. \qquad (2)$$

Here, (2) follows from Young's inequality and we have equality if and only if $\frac{|x_k|^p}{||\mathbf{x}||_p^p} = \frac{|y_k|^q}{||\mathbf{y}||_q^q}$. Summing each inequality in (2) for $1 \leq k \leq n$ yields

$$\begin{aligned} \frac{1}{||\mathbf{x}||_p\,||\mathbf{y}||_q}\sum_{k=1}^{n}|x_k y_k| &\leq \frac{1}{p}\cdot\frac{1}{||\mathbf{x}||_p^p}\underbrace{\left(\sum_{k=1}^{n}|x_k|^p\right)}_{||\mathbf{x}||_p^p} + \frac{1}{q}\cdot\frac{1}{||\mathbf{y}||_q^q}\underbrace{\left(\sum_{k=1}^{n}|y_k|^q\right)}_{||\mathbf{y}||_q^q}, \qquad (3) \\ &= \frac{1}{p} + \frac{1}{q} = 1. \qquad (4) \end{aligned}$$

Here, we have equality in (3) if and only if $\frac{|x_k|^p}{||\mathbf{x}||_p^p} = \frac{|y_k|^q}{||\mathbf{y}||_q^q}$ for all $1 \leq k \leq n$. Rearranging (4), we get

$$\sum_{k=1}^{n}|x_k y_k| \leq ||\mathbf{x}||_p\,||\mathbf{y}||_q\,,$$

which completes the proof.

(c) We will first prove the inequality mentioned in the hint. For this, let us first suppose $c = 0$. In this case, the inequality trivially holds with equality if and only if $c = Kd$ with $K = 0$. Similarly, if $d = 0$, the inequality trivially holds with equality if and only if $d = Kc$ with $K = 0$. Now suppose both $c, d \neq 0$. We can express c and d in polar form as $c = \rho_c e^{j\phi_c}$ and $d = \rho_d e^{j\phi_d}$, where $\rho_c, \rho_d > 0$ and $0 \leq \phi_c, \phi_d < 2\pi$. Then we have the following:

$$\begin{aligned} |c+d| &= \sqrt{(c+d)(c+d)^*} = \sqrt{\rho_c^2 + 2\rho_c\rho_d \cos(\phi_c - \phi_d) + \rho_d^2}\,, \\ &\leq \sqrt{\rho_c^2 + 2\rho_c\rho_d + \rho_d^2} = \sqrt{(\rho_c + \rho_d)^2} = \rho_c + \rho_d = |c| + |d|\,. \end{aligned} \tag{5}$$

Here, (5) follows from the fact that $\cos(x) \leq 1$ with equality if and only if $x \equiv 0 \pmod{2\pi}$ (i.e., $x = 2\pi m$ for some $m \in \mathbb{Z}$). As $-2\pi < \phi_c - \phi_d < 2\pi$, we have equality in (5) if and only if $\phi_c - \phi_d = 0$, or equivalently if and only if $\phi_c = \phi_d$. In this case, we either have $d = \rho_d e^{j\phi_c} = \frac{\rho_d}{\rho_c}\rho_c e^{j\phi_c} = Kc$, where $K = \frac{\rho_d}{\rho_c} > 0$ or $c = \rho_c e^{j\phi_d} = \frac{\rho_c}{\rho_d}\rho_d e^{j\phi_d} = Kd$, where $K = \frac{\rho_c}{\rho_d} > 0$. Hence, for $c, d \neq 0$, we have $|c+d| \leq |c| + |d|$ with equality if and only if $c = Kd$ or $d = Kc$ for some $K > 0$. Combining everything proven here, we have shown that $|c+d| \leq |c| + |d|$ with equality if and only if $c = Kd$ or $d = Kc$ for some $K \geq 0$.

Now to prove Minkowski's inequality, let us first consider the special case of $p = 1$. Using the triangle inequality proven above, it is clear that

$$|x_k + y_k| \leq |x_k| + |y_k|\,, \tag{6}$$

with equality if and only if $x_k = \lambda_k y_k$ or $y_k = \lambda_k x_k$ for some $\lambda_k \geq 0$. Summing the inequality in (6) over all k yields

$$||\mathbf{x} + \mathbf{y}||_1 \leq ||\mathbf{x}||_1 + ||\mathbf{y}||_1\,,$$

with equality if and only if $x_k = \lambda_k y_k$ or $y_k = \lambda_k x_k$ for some $\lambda_k \geq 0$ for all $1 \leq k \leq n$. This proves Minkowski's inequality for the special case of $p = 1$.
For $p > 1$, let us express Hölder's inequality in the following form for vectors $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$:

$$\sum_{k=1}^{n} |a_k b_k| \leq ||\mathbf{a}||_p \, ||\mathbf{b}||_{\frac{p}{p-1}}\,, \tag{7}$$

where $a_k \triangleq [\mathbf{a}]_k$ and $b_k \triangleq [\mathbf{b}]_k$ for $1 \leq k \leq n$. This follows since $\frac{1}{p} + \frac{1}{q} = 1$ is equivalent to $q = \frac{p}{p-1}$. We have equality in (7) if and only if $\frac{|a_k|^p}{||\mathbf{a}||_p} = \frac{|b_k|^{\frac{p}{p-1}}}{||\mathbf{b}||_{\frac{p}{p-1}}^{\frac{p}{p-1}}}$.
Setting $a_k = x_k$ and $b_k = (|x_k| + |y_k|)^{p-1}$ in (7) yields

$$\sum_{k=1}^{n} |x_k| \left(|x_k| + |y_k|\right)^{p-1} \leq ||\mathbf{x}||_p \left(\sum_{k=1}^{n} \left(|x_k| + |y_k|\right)^p\right)^{\frac{p-1}{p}}\,. \tag{8}$$

Similarly, setting $a_k = y_k$ and $b_k = (|x_k| + |y_k|)^{p-1}$ in (7) yields

$$\sum_{k=1}^{n} |y_k| \left(|x_k| + |y_k|\right)^{p-1} \leq ||\mathbf{y}||_p \left(\sum_{k=1}^{n} \left(|x_k| + |y_k|\right)^p\right)^{\frac{p-1}{p}}\,. \tag{9}$$

Adding (8) and (9) yields the following:

$$\sum_{k=1}^{n} (|x_k| + |y_k|)^p \leq \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{p-1}{p}} \left(||\mathbf{x}||_p + ||\mathbf{y}||_p \right) .$$

Dividing both sides of the above equation by the common term yields

$$\left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{1}{p}} \leq ||\mathbf{x}||_p + ||\mathbf{y}||_p . \quad (10)$$

Tracing back the conditions for equality from (8) and (9) in (10), it follows that we have equality in (10) if and only if $\frac{|x_k|^p}{||\mathbf{x}||_p} = \frac{|y_k|^p}{||\mathbf{y}||_p}$ for all $1 \leq k \leq n$. Equivalently, we have equality in (10) if and only if $|x_k| = C\,|y_k|$ or $|y_k| = C\,|x_k|$ for some $C > 0$ for all $1 \leq k \leq n$. By the triangle inequality for complex scalars, though, we have

$$\left(\sum_{k=1}^{n} (|x_k + y_k|)^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{1}{p}} , \quad (11)$$

with equality if and only if $x_k = \lambda_k y_k$ or $y_k = \lambda_k x_k$ for some $\lambda_k \geq 0$ for all $1 \leq k \leq n$. Combining (10) and (11), along with their respective conditions for equality, it follows that

$$\left(\sum_{k=1}^{n} (|x_k + y_k|)^p \right)^{\frac{1}{p}} \leq ||\mathbf{x}||_p + ||\mathbf{y}||_p ,$$

with equality if and only if $\mathbf{x} = \lambda \mathbf{y}$ or $\mathbf{y} = \lambda \mathbf{x}$ for some $\lambda \geq 0$. This completes the proof of Minkowski's inequality for $p > 1$.

3. For $\mathbf{z} = \mathbf{0}$, we clearly have $||\mathbf{z}||_{p\circledast} = ||\mathbf{z}||_q = \mathbf{0}$. To prove the result for $\mathbf{z} \neq \mathbf{0}$, for convenience of notation, set $x_k = [\mathbf{x}]_k$ and $z_k = [\mathbf{z}]_k$ for $1 \leq k \leq n$. Then, note that we have

$$\left|\mathbf{z}^\dagger \mathbf{x}\right| = \left| \sum_{k=1}^{n} z_k^* x_k \right| \leq \sum_{k=1}^{n} |z_k^* x_k| = \sum_{k=1}^{n} |x_k z_k| \leq ||\mathbf{x}||_p \, ||\mathbf{z}||_q = ||\mathbf{z}||_q , \quad (12)$$

where the last inequality follows from Hölder's inequality. If we set

$$x_k = \frac{|z_k|^{q-2} \, z_k}{||\mathbf{z}||_q^{q-1}} , \quad (13)$$

then we have

$$\left|\mathbf{z}^\dagger \mathbf{x}\right| = \left| \sum_{k=1}^{n} z_k^* \frac{|z_k|^{q-2} \, z_k}{||\mathbf{z}||_q^{q-1}} \right| = \frac{1}{||\mathbf{z}||_q^{q-1}} \left(\sum_{k=1}^{n} |z_k|^q \right) = \frac{1}{||\mathbf{z}||_q^{q-1}} \cdot ||\mathbf{z}||_q^q = ||\mathbf{z}||_q ,$$

and so we have equality in (12). However, from (13), we also have

$$\begin{aligned} ||\mathbf{x}||_p &= \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} = \left(\sum_{k=1}^{n} \frac{|z_k|^{(q-1)p}}{||\mathbf{z}||_q^{(q-1)p}} \right)^{\frac{1}{p}} = \left(\frac{1}{||\mathbf{z}||_q^q} \sum_{k=1}^{n} |z_k|^q \right)^{\frac{1}{p}} , \\ &= \left(\frac{1}{||\mathbf{z}||_q^q} \cdot ||\mathbf{z}||_q^q \right)^{\frac{1}{p}} = (1)^{\frac{1}{p}} = 1 . \end{aligned} \quad (14)$$

Here, we used the fact that $(q-1)\,p = q$ in (14), which follows from the fact that $\frac{1}{p}+\frac{1}{q}=1$. Thus, the choice of $\mathbf{x}$ given in (13) satisfies the condition that $||\mathbf{x}||_p = 1$ and achieves the upper bound given in (12). From this, we conclude that

$$||\mathbf{z}||_{p^\circledast} = ||\mathbf{z}||_q = ||\mathbf{z}||_{\frac{p}{p-1}} \,.$$

4. (a) Note that we have

$$\begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix}.$$

Similarly, we have

$$\begin{bmatrix} \mathbf{I}_m & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_n \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{0}_{m\times n} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{D} \end{bmatrix}.$$

(b) The first part follows directly from applying the first set of hints given in this subproblem to the results of part (a). To prove the second part, we use the fact that

$$\det\left(\begin{bmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{0}_{n\times n} & \mathbf{I}_n \end{bmatrix}\right) = \det\left(\begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{Y} & \mathbf{I}_n \end{bmatrix}\right) = \prod_{k=1}^{m+n} 1 = 1\,,$$

for any $m \times n$ matrix $\mathbf{X}$ and $n \times m$ matrix $\mathbf{Y}$ along with the second set of hints given in this subproblem to the results of the first part of this subproblem.

(c) This follows directly from the hints given in this subproblem along with the hints given in the previous subproblem.

(d) Using the results from part (c), we have

$$\begin{aligned}
\mathbf{M}^{-1} &= \begin{bmatrix} \mathbf{I}_m & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{bmatrix}, \\
&= \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1} \\ \mathbf{0}_{n\times m} & \left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{bmatrix}, \\
&= \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1} \\ -\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}^{-1} & \left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1} \end{bmatrix}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathbf{M}^{-1} &= \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix}, \\
&= \begin{bmatrix} \left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{0}_{m\times n} \\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{bmatrix}, \\
&= \begin{bmatrix} \left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & -\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1} + \mathbf{D}^{-1} \end{bmatrix}.
\end{aligned}$$

Comparing the $(1,1)$-th block of both expressions for $\mathbf{M}^{-1}$ yields

$$\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}=\mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}^{-1},$$

as desired.

***5.** For convenience, define the following vectors:

$$\overline{\boldsymbol{v}} \triangleq \boldsymbol{v}-\boldsymbol{\mu}_{\mathbf{x}}\,,\ \overline{\boldsymbol{y}} \triangleq \boldsymbol{y}_0-\boldsymbol{\mu}_{\mathbf{y}}\,,\ \overline{\boldsymbol{z}} \triangleq \begin{bmatrix} \overline{\boldsymbol{v}} \\ \overline{\boldsymbol{y}} \end{bmatrix}. \tag{15}$$

Then, from the hint given in the problem, we have

$$\begin{aligned} f_{\mathbf{v}}(\boldsymbol{v}) &= \frac{\dfrac{1}{(2\pi)^{\frac{m+n}{2}}(\det(\boldsymbol{\Sigma}_{\mathbf{z}}))^{\frac{1}{2}}}e^{-\frac{1}{2}\overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\boldsymbol{z}}}{\dfrac{1}{(2\pi)^{\frac{n}{2}}(\det(\boldsymbol{\Sigma}_{\mathbf{y}}))^{\frac{1}{2}}}e^{-\frac{1}{2}\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{y}}}, \\ &= \frac{1}{(2\pi)^{\frac{m}{2}}}\left(\frac{\det(\boldsymbol{\Sigma}_{\mathbf{y}})}{\det(\boldsymbol{\Sigma}_{\mathbf{z}})}\right)^{\frac{1}{2}} e^{-\frac{1}{2}\left[\overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\overline{\boldsymbol{z}}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}\right]}. \end{aligned} \tag{16}$$

From the results from part (b) of the previous problem, we have

$$\det(\boldsymbol{\Sigma}_{\mathbf{z}})=\det(\boldsymbol{\Sigma}_{\mathbf{v}})\det(\boldsymbol{\Sigma}_{\mathbf{y}})\ .$$

Using this in (16), we get

$$f_{\mathbf{v}}(\boldsymbol{v})=\frac{1}{(2\pi)^{\frac{m}{2}}\left(\det(\boldsymbol{\Sigma}_{\mathbf{v}})\right)^{\frac{1}{2}}}e^{-\frac{1}{2}\left[\overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\overline{\boldsymbol{z}}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}\right]}. \tag{17}$$

Now, from the results of part (d) of the previous problem, we have

$$\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}=\begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} & -\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \\ -\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1} & \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}+\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \end{bmatrix}.$$

Using this, we have the following:

$$\begin{aligned} \overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\overline{\boldsymbol{z}} &= \begin{bmatrix} \overline{\boldsymbol{v}}^T & \overline{\boldsymbol{y}}^T \end{bmatrix}\begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} & -\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \\ -\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1} & \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}+\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \end{bmatrix}\begin{bmatrix} \overline{\boldsymbol{v}} \\ \overline{\boldsymbol{y}} \end{bmatrix}, \\ &= \begin{bmatrix} \left(\overline{\boldsymbol{v}}^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\right) & \left(-\overline{\boldsymbol{v}}^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}+\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}+\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\right) \end{bmatrix}\begin{bmatrix} \overline{\boldsymbol{v}} \\ \overline{\boldsymbol{y}} \end{bmatrix}, \\ &= \overline{\boldsymbol{v}}^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\overline{\boldsymbol{v}}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\overline{\boldsymbol{v}}-\overline{\boldsymbol{v}}^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}+\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}+\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}. \end{aligned}$$

From this, it can be seen that we have the following:

$$\overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\overline{\boldsymbol{z}}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}=\left(\overline{\boldsymbol{v}}-\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}\right)^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\left(\overline{\boldsymbol{v}}-\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}\right), \tag{18}$$

where we used the fact that $\boldsymbol{\Sigma}_{\mathbf{yx}}=\boldsymbol{\Sigma}_{\mathbf{xy}}^T$. Using (15) in (18), it is clear that we have

$$\overline{\boldsymbol{z}}^T\boldsymbol{\Sigma}_{\mathbf{z}}^{-1}\overline{\boldsymbol{z}}-\overline{\boldsymbol{y}}^T\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\overline{\boldsymbol{y}}=\left(\boldsymbol{v}-\boldsymbol{\mu}_{\mathbf{v}}\right)^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\left(\boldsymbol{v}-\boldsymbol{\mu}_{\mathbf{v}}\right),$$

which, upon substitution into (17), leads to the following:

$$f_{\mathbf{v}}(\boldsymbol{v})=\frac{1}{(2\pi)^{\frac{m}{2}}\left(\det(\boldsymbol{\Sigma}_{\mathbf{v}})\right)^{\frac{1}{2}}}e^{-\frac{1}{2}(\boldsymbol{v}-\boldsymbol{\mu}_{\mathbf{v}})^T\boldsymbol{\Sigma}_{\mathbf{v}}^{-1}(\boldsymbol{v}-\boldsymbol{\mu}_{\mathbf{v}})}.$$

This shows that $\mathbf{v}\sim\mathcal{N}(\boldsymbol{\mu}_{\mathbf{v}},\boldsymbol{\Sigma}_{\mathbf{v}})$ as desired.