香港中文大學
The Chinese University of Hong Kong

11.9 Zero-Mean Gaussian Noise is the Worst Additive Noise

Worst Additive Noise

Worst Additive Noise

- We will show that in terms of the capacity of the system, the zero-mean Gaussian noise is the worst additive noise given that the noise vector has a fixed correlation matrix.

Worst Additive Noise

- We will show that in terms of the capacity of the system, the zero-mean Gaussian noise is the worst additive noise given that the noise vector has a fixed correlation matrix.

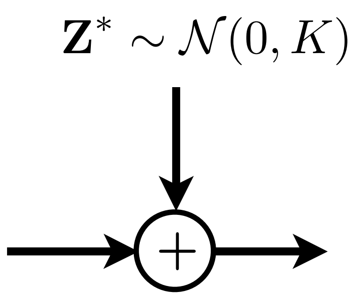
- The diagonal elements of the correlation matrix specify the power of the individual noise variables.

Worst Additive Noise

- We will show that in terms of the capacity of the system, the zero-mean Gaussian noise is the worst additive noise given that the noise vector has a fixed correlation matrix.
- The diagonal elements of the correlation matrix specify the power of the individual noise variables.
- The other elements in the matrix give a characterization of the correlation between the noise variables.

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

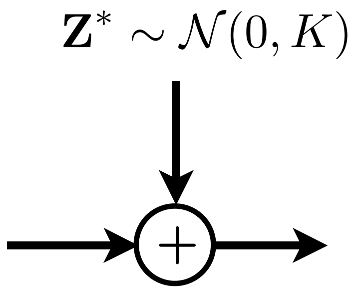
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

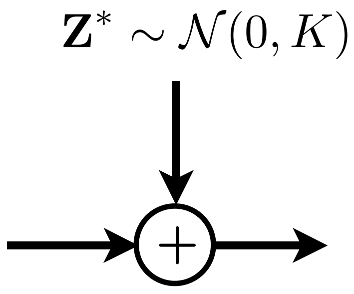
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

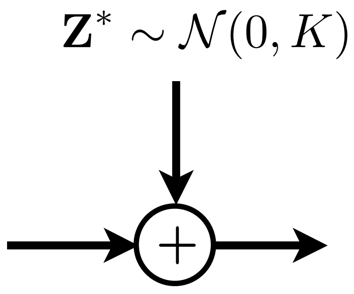
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

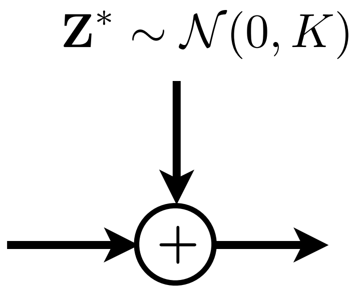
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

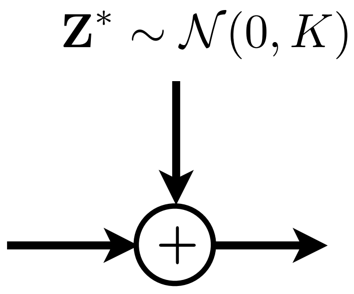
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

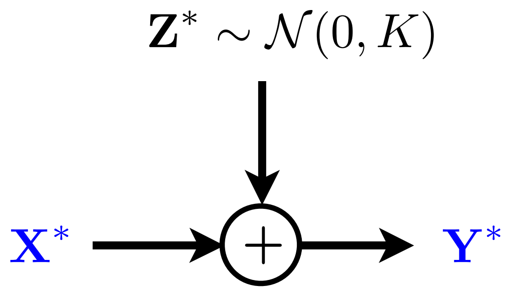
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

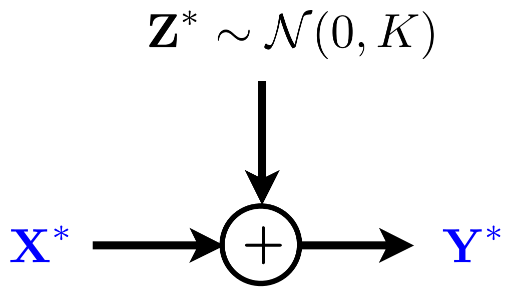
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

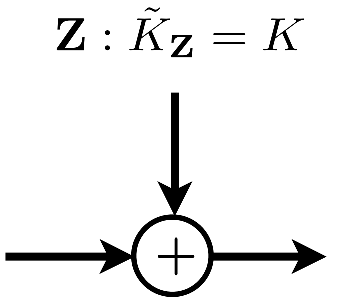
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

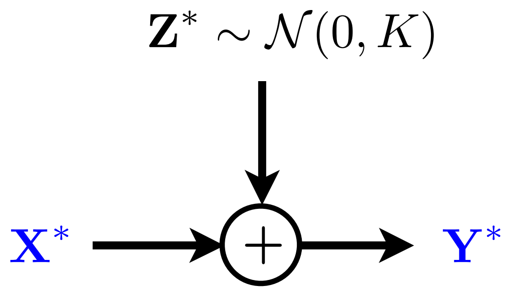
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

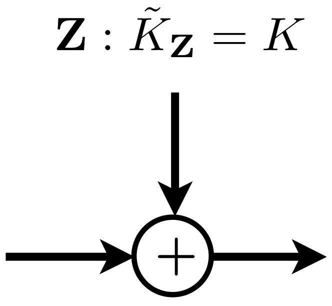
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

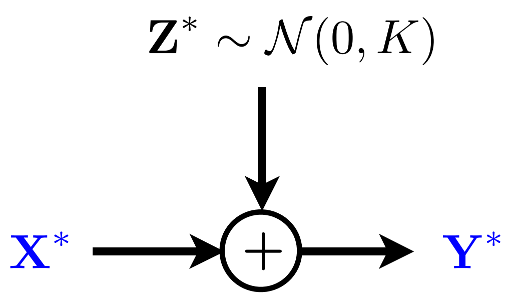
Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

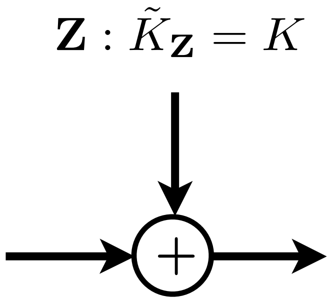
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System

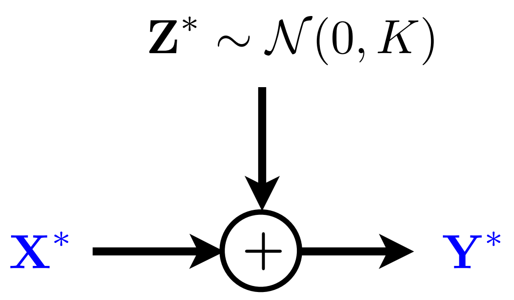


1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

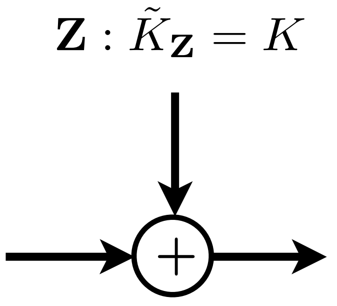
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



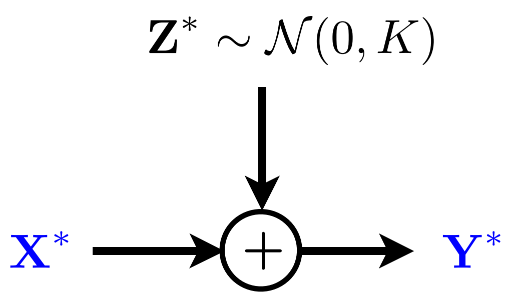
1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

3. Let C be the capacity of the system.

Main Idea

Zero-Mean Gaussian System



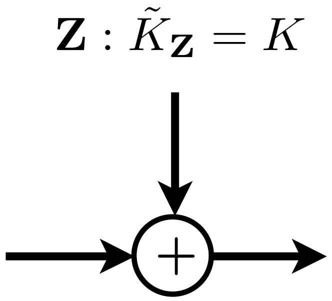
1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

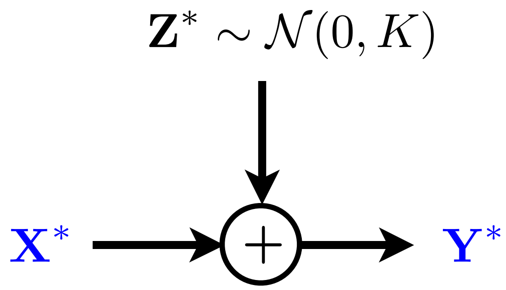
2. Assume that the joint pdf of $\mathbf{Z}$ exists.

3. Let C be the capacity of the system.

4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

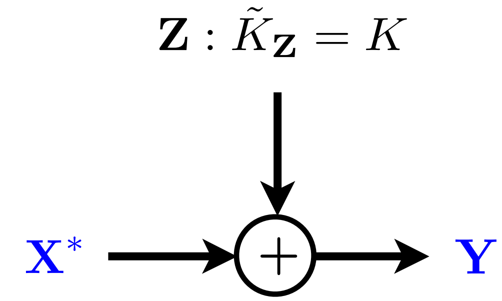
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

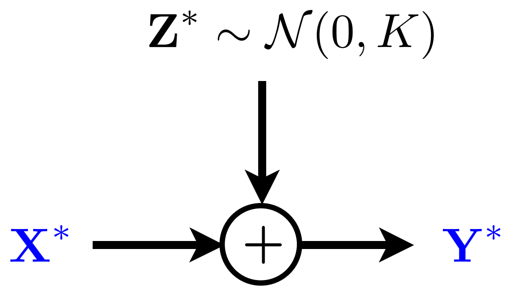
3. Let C be the capacity of the system.

4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

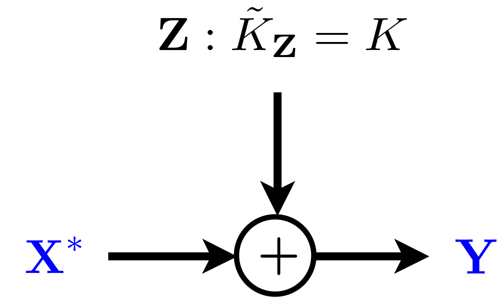
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

3. Let C be the capacity of the system.

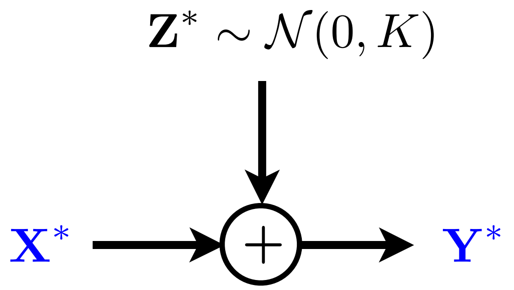
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.

Main Idea

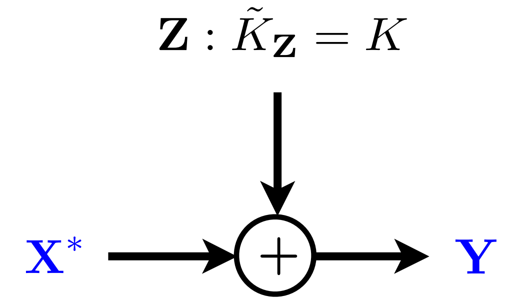
Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.
2. Let C^* be the capacity of the system.
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.
4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



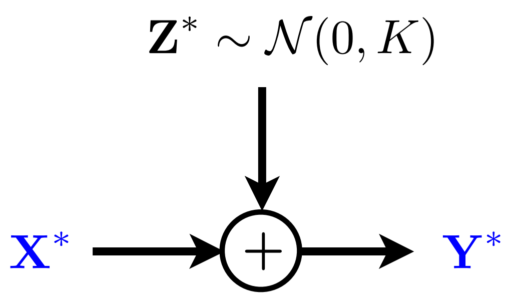
1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.
2. Assume that the joint pdf of $\mathbf{Z}$ exists.
3. Let C be the capacity of the system.
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \leq C$.

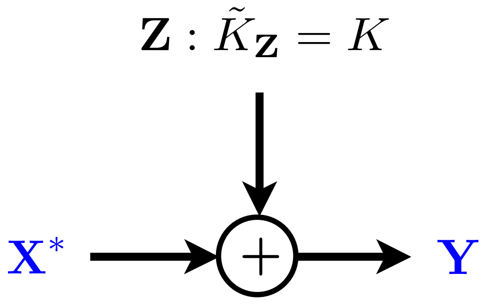
Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.
2. Let C^* be the capacity of the system.
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.
4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System

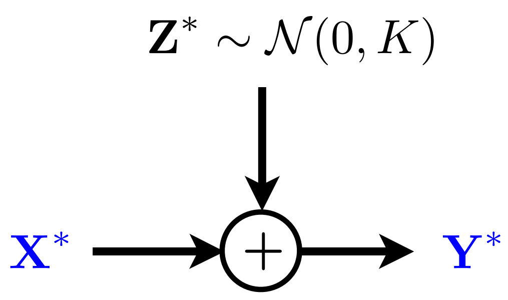


1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.
2. Assume that the joint pdf of $\mathbf{Z}$ exists.
3. Let C be the capacity of the system.
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \le C$.
- We will show that $I(\mathbf{X}^*; \mathbf{Y}^*) \le I(\mathbf{X}^*; \mathbf{Y})$.

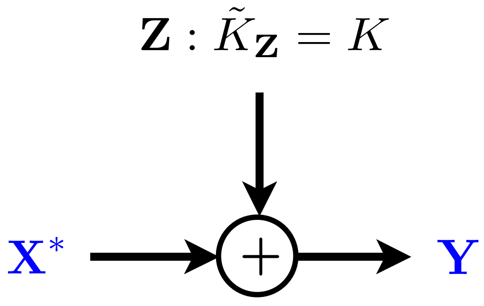
Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.
2. Let C^* be the capacity of the system.
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.
4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System

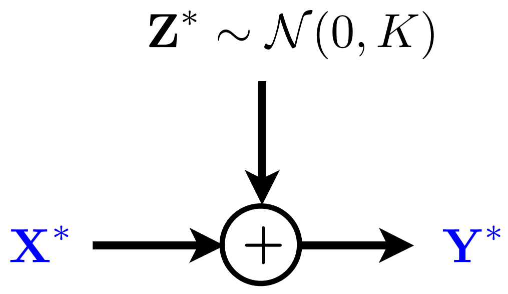


1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.
2. Assume that the joint pdf of $\mathbf{Z}$ exists.
3. Let C be the capacity of the system.
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \le C$.
- We will show that $I(\mathbf{X}^*; \mathbf{Y}^*) \le I(\mathbf{X}^*; \mathbf{Y})$.
- Hence,
$$C^* = I(\mathbf{X}^*; \mathbf{Y}^*) \le I(\mathbf{X}^*; \mathbf{Y}) \le C.$$

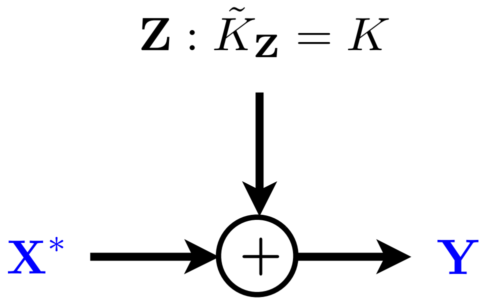
Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.
2. Let C^* be the capacity of the system.
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.
4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System

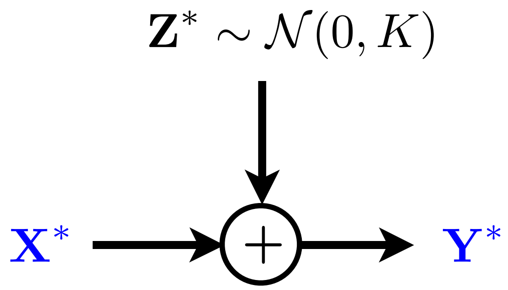


1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.
2. Assume that the joint pdf of $\mathbf{Z}$ exists.
3. Let C be the capacity of the system.
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \leq C$.
- We will show that $I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y})$.
- Hence,
$$C^* = I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y}) \leq C.$$

Main Idea

Zero-Mean Gaussian System



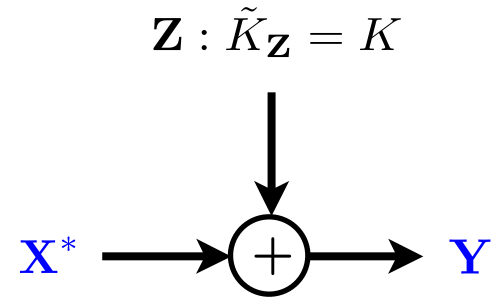
1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

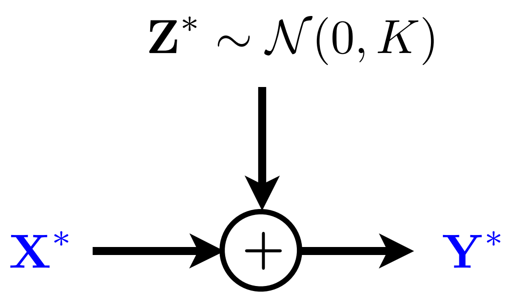
3. Let C be the capacity of the system.

4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \leq C$.
- We will show that $\underline{I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y})}$.
- Hence,
$$C^* = \underline{I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y})} \leq C.$$

Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.

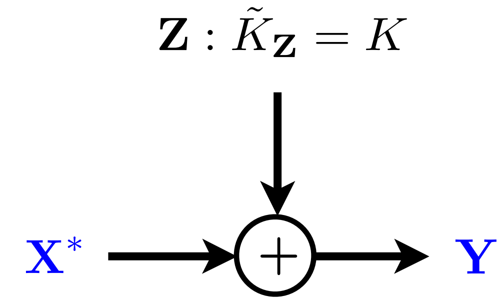
2. Let C^* be the capacity of the system.

3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.

4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.

2. Assume that the joint pdf of $\mathbf{Z}$ exists.

3. Let C be the capacity of the system.

4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,

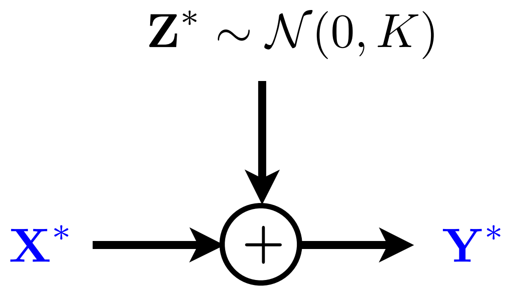
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $\underline{I(\mathbf{X}^*; \mathbf{Y}) \leq C}$.
- We will show that $I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y})$.
- Hence,

$$C^* = I(\mathbf{X}^*; \mathbf{Y}^*) \leq \underline{I(\mathbf{X}^*; \mathbf{Y}) \leq C}.$$

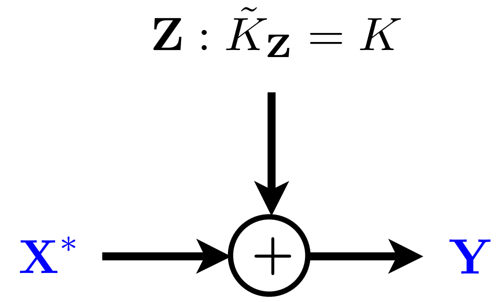
Main Idea

Zero-Mean Gaussian System



1. Consider a system of correlated Gaussian channels with noise vector $\mathbf{Z}^* \sim \mathcal{N}(0, K)$, and so $\tilde{K}_{\mathbf{Z}^*} = K$.
2. Let C^* be the capacity of the system.
3. Let $\mathbf{X}^*$ be the zero-mean Gaussian input vector that achieves the capacity.
4. Let $\mathbf{Y}^*$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y}^* = \mathbf{X}^* + \mathbf{Z}^*.$$

Alternative System



1. Consider a system exactly the same as the Zero-Mean Gaussian System except that the noise vector $\mathbf{Z}$, which has the same correlation matrix as $\mathbf{Z}^*$, may neither be zero-mean nor Gaussian.
2. Assume that the joint pdf of $\mathbf{Z}$ exists.
3. Let C be the capacity of the system.
4. Let $\mathbf{Y}$ be the output of the system with $\mathbf{X}^*$ as input, i.e.,
$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z}.$$

- For the Zero-Mean Gaussian System, $C^* = I(\mathbf{X}^*; \mathbf{Y}^*)$.
- For the alternative system, $I(\mathbf{X}^*; \mathbf{Y}) \leq C$.
- We will show that $I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y})$.
- Hence,
$$C^* = I(\mathbf{X}^*; \mathbf{Y}^*) \leq I(\mathbf{X}^*; \mathbf{Y}) \leq C.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Lemma 11.33 Let $\mathbf{X}$ be a zero-mean random vector and

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z}$ is independent of $\mathbf{X}$. Then

$$\tilde{K}_{\mathbf{Y}} = \tilde{K}_{\mathbf{X}} + \tilde{K}_{\mathbf{Z}}.$$

Lemma 11.33 Let $\mathbf{X}$ be a zero-mean random vector and

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z}$ is independent of $\mathbf{X}$. Then

$$\tilde{K}_{\mathbf{Y}} = \tilde{K}_{\mathbf{X}} + \tilde{K}_{\mathbf{Z}}.$$

Remark The scalar case has been proved in the proof of Theorem 11.21.

Lemma 11.33 Let $\mathbf{X}$ be a zero-mean random vector and

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z}$ is independent of $\mathbf{X}$. Then

$$\tilde{K}_{\mathbf{Y}} = \tilde{K}_{\mathbf{X}} + \tilde{K}_{\mathbf{Z}}.$$

Remark The scalar case has been proved in the proof of Theorem 11.21.

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Lemma 11.33 Let $\mathbf{X}$ be a zero-mean random vector and

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z}$ is independent of $\mathbf{X}$. Then

$$\tilde{K}_{\mathbf{Y}} = \tilde{K}_{\mathbf{X}} + \tilde{K}_{\mathbf{Z}}.$$

Remark The scalar case has been proved in the proof of Theorem 11.21.

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Remark A similar technique has been used in proving Theorems 2.50 and 10.41 (maximum entropy distributions).

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$H(p^*) - H(p)$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \tag{1}$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ = & \quad - \sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \end{aligned}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^m \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ &= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \end{aligned}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ &= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \end{aligned}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned}
& H(p^*) - H(p) \\
&= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\
&= D(p \| p^*)
\end{aligned}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \le i \le m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ &= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ &= D(p \| p^*) \\ &\ge 0. \end{aligned}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \le i \le m. \tag{1}$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned}
& H(p^*) - H(p) \\
&= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\
&= D(p \| p^*) \\
&\ge 0.
\end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \le i \le m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned}
& H(p^*) - H(p) \\
&= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\
&= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\
&= D(p \| p^*) \\
&\ge 0.
\end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \tag{1}$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} H(p^*) - H(p) & \\ = & -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ = & D(p \| p^*) \\ \geq & 0. \end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 10.41 Let

$$f^*(\mathbf{x}) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(\mathbf{x})}$$

for all $\mathbf{x} \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\int_{\mathcal{S}_f} r_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = a_i \quad \text{for } 1 \leq i \leq m. \tag{2}$$

Then f^* maximizes $h(f)$ over all pdf f defined on $\mathcal{S}$, subject to the constraints in (2).

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \tag{1}$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} H(p^*) - H(p) \\ &= -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ &= \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ &= D(p \| p^*) \\ &\geq 0. \end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 10.41 Let

$$f^*(\mathbf{x}) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(\mathbf{x})}$$

for all $\mathbf{x} \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\int_{\mathcal{S}_f} r_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = a_i \quad \text{for } 1 \leq i \leq m. \tag{2}$$

Then f^* maximizes $h(f)$ over all pdf f defined on $\mathcal{S}$, subject to the constraints in (2).

Remark The key step is to establish that

$$\int_{\mathcal{S}} f^*(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{S}_f} f(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x}. \tag{3}$$

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \le i \le m. \tag{1}$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ = & -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ = & D(p \| p^*) \\ \ge & 0. \end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 10.41 Let

$$f^*(\mathbf{x}) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(\mathbf{x})}$$

for all $\mathbf{x} \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\int_{\mathcal{S}_f} r_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = a_i \quad \text{for } 1 \le i \le m. \tag{2}$$

Then f^* maximizes $h(f)$ over all pdf f defined on $\mathcal{S}$, subject to the constraints in (2).

Remark The key step is to establish that

$$\int_{\mathcal{S}} f^*(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{S}_f} f(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x}. \tag{3}$$

Theorem 10.45 Let $\mathbf{X}$ be a vector of n continuous random variables with correlation matrix $\tilde{K}$. Then

$$h(\mathbf{X}) \le \frac{1}{2} \log \left[(2\pi e)^n |\tilde{K}| \right]$$

with equality if and only if $\mathbf{X} \sim \mathcal{N}(0, \tilde{K})$.

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \leq i \leq m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ = & \quad -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & \quad -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & \quad \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ = & \quad D(p \| p^*) \\ \geq & \quad 0. \end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 10.41 Let

$$f^*(\mathbf{x}) = e^{-\lambda_0 - \sum_{i=1}^{m} \lambda_i r_i(\mathbf{x})}$$

for all $\mathbf{x} \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\int_{\mathcal{S}_f} r_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = a_i \quad \text{for } 1 \leq i \leq m. \qquad (2)$$

Then f^* maximizes $h(f)$ over all pdf f defined on $\mathcal{S}$, subject to the constraints in (2).

Remark The key step is to establish that

$$\int_{\mathcal{S}} f^*(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{S}_f} f(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x}. \qquad (3)$$

Theorem 10.45 Let $\mathbf{X}$ be a vector of n continuous random variables with correlation matrix $\tilde{K}$. Then

$$h(\mathbf{X}) \leq \frac{1}{2} \log \left[(2\pi e)^n |\tilde{K}| \right]$$

with equality if and only if $\mathbf{X} \sim \mathcal{N}(0, \tilde{K})$.

Then (3) and Theorem 10.45 together imply

Theorem 2.50 Let

$$p^*(x) = e^{-\lambda_0 - \sum_{i=1}^m \lambda_i r_i(x)}$$

for all $x \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\sum_{x \in \mathcal{S}_p} p(x) r_i(x) = a_i \quad \text{for } 1 \le i \le m. \qquad (1)$$

Then p^* maximizes $H(p)$ over all probability distribution p on $\mathcal{S}$ subject to (1).

Sketch of Proof

$$\begin{aligned} & H(p^*) - H(p) \\ = & -\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & -\sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x) + \sum_{x \in \mathcal{S}_p} p(x) \ln p(x) \\ = & \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{p^*(x)} \\ = & D(p \| p^*) \\ \ge & 0. \end{aligned}$$

Remark The key step is to establish that

$$\sum_{x \in \mathcal{S}} p^*(x) \ln p^*(x) = \sum_{x \in \mathcal{S}_p} p(x) \ln p^*(x).$$

Theorem 10.41 Let

$$f^*(\mathbf{x}) = e^{-\lambda_0 - \sum_{i=1}^m \lambda_i r_i(\mathbf{x})}$$

for all $\mathbf{x} \in \mathcal{S}$, where $\lambda_0, \lambda_1, \cdots, \lambda_m$ are chosen such that

$$\int_{\mathcal{S}_f} r_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = a_i \quad \text{for } 1 \le i \le m. \qquad (2)$$

Then f^* maximizes $h(f)$ over all pdf f defined on $\mathcal{S}$, subject to the constraints in (2).

Remark The key step is to establish that

$$\int_{\mathcal{S}} f^*(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{S}_f} f(\mathbf{x}) \ln f^*(\mathbf{x}) d\mathbf{x}. \qquad (3)$$

Theorem 10.45 Let $\mathbf{X}$ be a vector of n continuous random variables with correlation matrix $\tilde{K}$. Then

$$h(\mathbf{X}) \le \frac{1}{2} \log \left[(2\pi e)^n |\tilde{K}| \right]$$

with equality if and only if $\mathbf{X} \sim \mathcal{N}(0, \tilde{K})$.

Then (3) and Theorem 10.45 together imply

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Data Processing Inequality for Informational Divergence

Theorem Let X, X', Y, and Y' be real random variables such that f_{XY} and $f_{X'Y'}$ exist, and $f_{Y|X} = f_{Y'|X'}$. Then

$$D(f_X \| f_{X'}) \geq D(f_Y \| f_{Y'}).$$

Data Processing Inequality for Informational Divergence

Theorem Let X, X', Y, and Y' be real random variables such that f_{XY} and $f_{X'Y'}$ exist, and $f_{Y|X} = f_{Y'|X'}$. Then

$$D(f_X \| f_{X'}) \geq D(f_Y \| f_{Y'}).$$

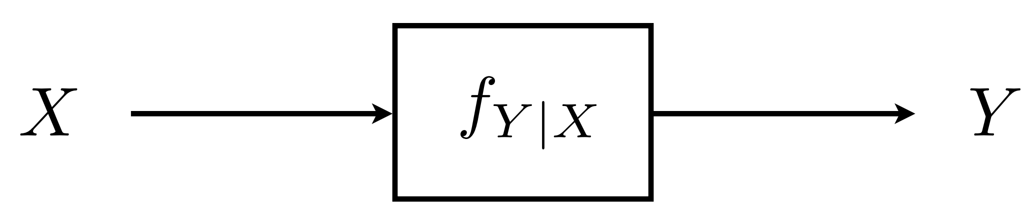
Proof Exercise.

Data Processing Inequality for Informational Divergence

Theorem Let X, X', Y, and Y' be real random variables such that f_{XY} and $f_{X'Y'}$ exist, and $f_{Y|X} = f_{Y'|X'}$. Then

$$D(f_X \| f_{X'}) \geq D(f_Y \| f_{Y'}).$$

Proof Exercise.

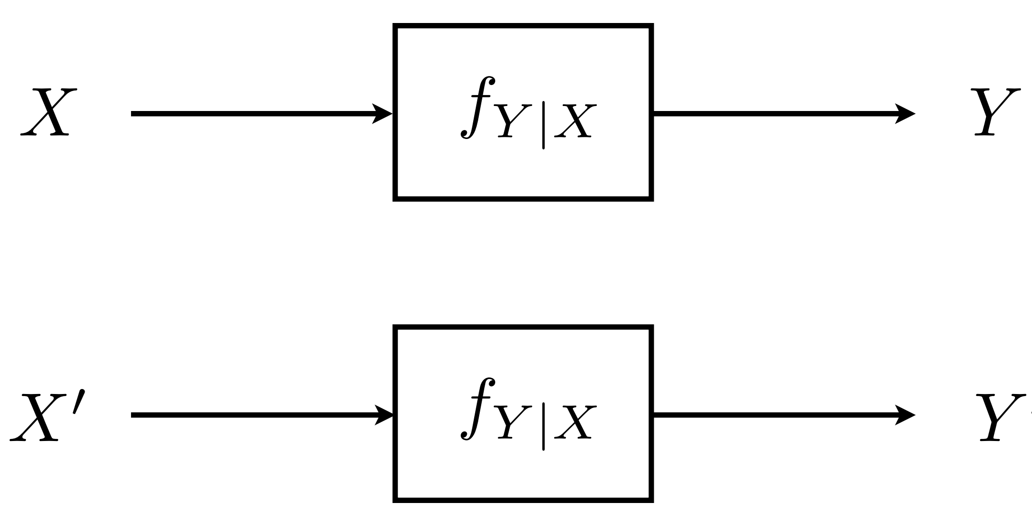


Data Processing Inequality for Informational Divergence

Theorem Let X, X', Y, and Y' be real random variables such that f_{XY} and $f_{X'Y'}$ exist, and $f_{Y|X} = f_{Y'|X'}$. Then

$$D(f_X \| f_{X'}) \geq D(f_Y \| f_{Y'}).$$

Proof Exercise.



Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

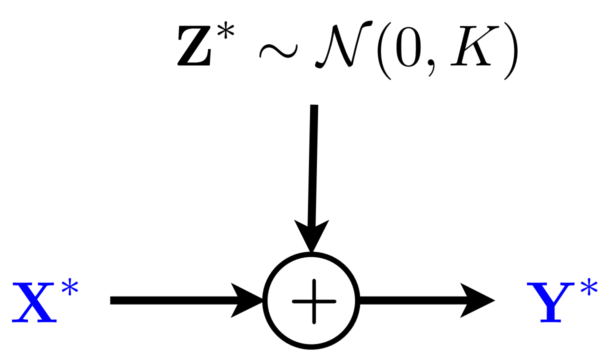
where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

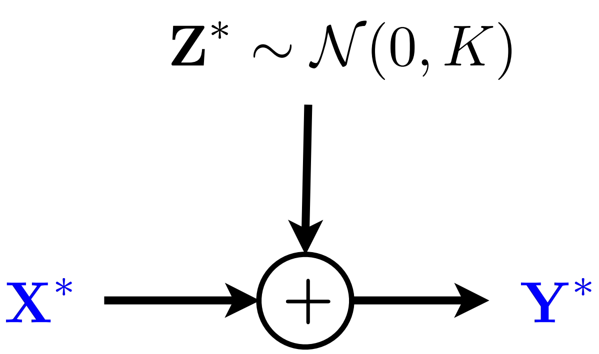


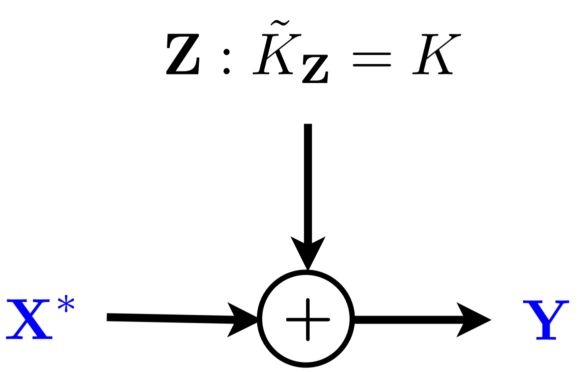
Proof

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.



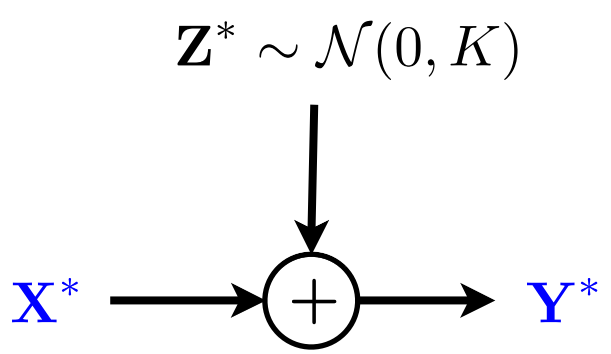


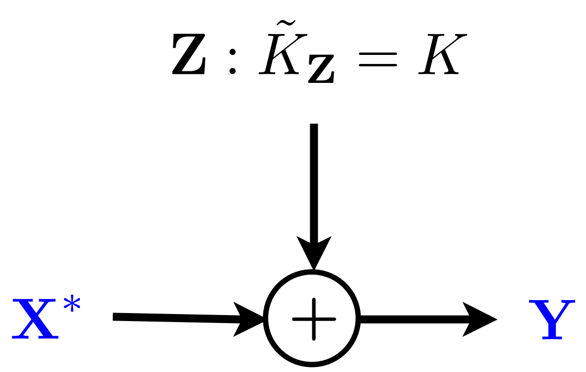
Proof

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





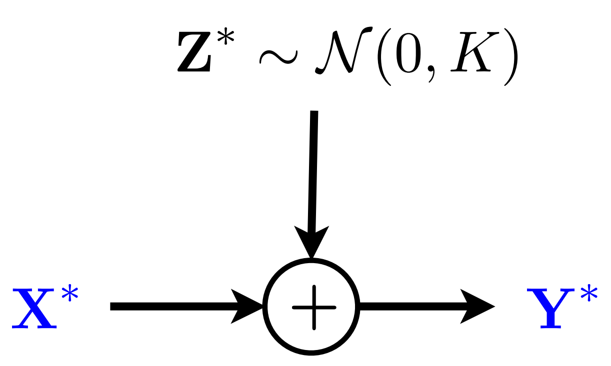
Proof

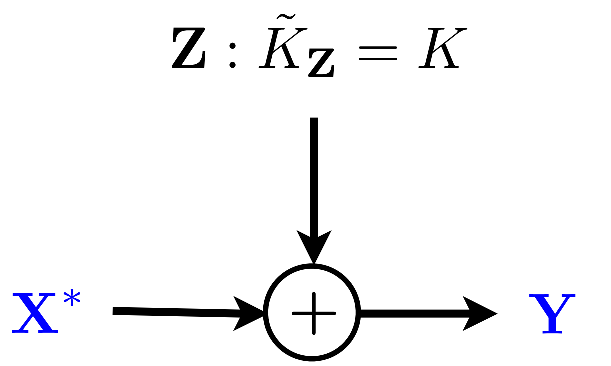
1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

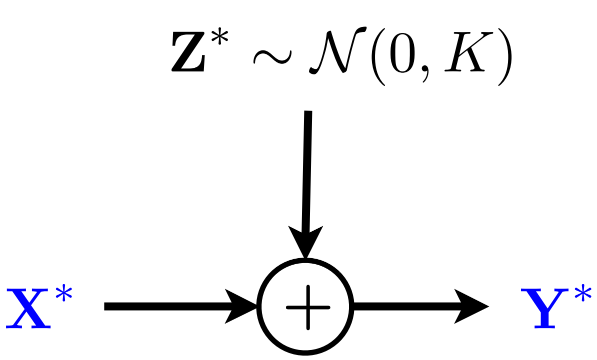
1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

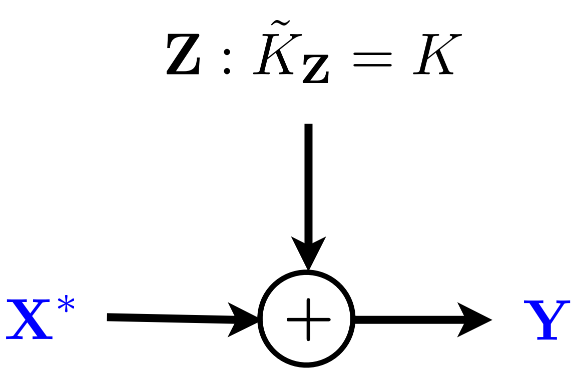
2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

Lemma 11.33 Let $\mathbf{X}$ be a zero-mean random vector and

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z}$ is independent of $\mathbf{X}$. Then

$$\tilde{K}_{\mathbf{Y}} = \tilde{K}_{\mathbf{X}} + \tilde{K}_{\mathbf{Z}}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y})$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\ = \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\ = \quad & \underline{h(\mathbf{Y}^*)} - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&\quad = \quad h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&\quad = \quad -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad\quad\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) \underline{- h(\mathbf{Z}^*)} - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&\quad = \quad h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&\quad = \quad -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\qquad\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&\quad = h(\mathbf{Y}^*) - h(\mathbf{Z}^*) \underline{- h(\mathbf{Y})} + h(\mathbf{Z}) \\
&\quad = -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\qquad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*; \mathbf{Y}^*) - I(\mathbf{X}^*; \mathbf{Y}) \\ = \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) \underline{- h(\mathbf{Y})} + h(\mathbf{Z}) \\ = \quad & -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\ & \underline{+ \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*; \mathbf{Y}^*) - I(\mathbf{X}^*; \mathbf{Y}) \\ = \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + \underline{h(\mathbf{Z})} \\ = \quad & -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\ & + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&\quad= \quad h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + \underline{h(\mathbf{Z})} \\
&\quad= \quad -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\qquad\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} \underline{- \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*; \mathbf{Y}^*) - I(\mathbf{X}^*; \mathbf{Y}) \\ &= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\ &= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\ &\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \end{aligned}$$

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned} & I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\ = \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\ = \quad & -\int \underline{f_{\mathbf{Y}^*}(\mathbf{y})} \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\ & + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\ = \quad & -\int \underline{f_{\mathbf{Y}}(\mathbf{y})} \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\ & + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \end{aligned}$$

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*; \mathbf{Y}^*) - I(\mathbf{X}^*; \mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int \underline{f_{\mathbf{Z}^*}(\mathbf{z})} \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int \underline{f_{\mathbf{Z}^*}(\mathbf{z})} \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} \underline{f_{\mathbf{Z}}(\mathbf{z})} \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Lemma 11.34 Let $\mathbf{Y}^* \sim \mathcal{N}(0, K)$ and $\mathbf{Y}$ be any random vector with correlation matrix K. Then

$$\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{S}_{\mathbf{Y}}} f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}.$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
= \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
= \quad & -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
& + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
= \quad & \underline{-\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
& + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \underline{-\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y}} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \underline{\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y})d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z})d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y} - \int_{\mathcal{S}_\mathbf{Z}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z} \\
&= \underline{-\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y})d\mathbf{y}} + \int_{\mathcal{S}_\mathbf{Z}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z})d\mathbf{z} \\
&\quad + \underline{\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y}} - \int_{\mathcal{S}_\mathbf{Z}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z} \\
&= \underline{\int \log\left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})}\right) f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y}} - \int_{\mathcal{S}_\mathbf{Z}} \log\left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})}\right) f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \underline{\int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z}} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \underline{\int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z}} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} \underline{- \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y})d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z})d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y})d\mathbf{y} + \underline{\int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z})d\mathbf{z}} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y} \underline{- \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z}} \\
&= \int \log\left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})}\right) f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y} - \underline{\int_{\mathcal{S}_{\mathbf{Z}}} \log\left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})}\right) f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z}}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
= \quad & h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
= \quad & -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
& + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
= \quad & -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
& + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
= \quad & \underline{\int \log\left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})}\right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}} - \int_{\mathcal{S}_{\mathbf{Z}}} \log\left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})}\right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log\left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})}\right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log\left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})}\right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*; \mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*; \mathbf{Y}^*) - I(\mathbf{X}^*; \mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \underline{\int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - \underline{D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})}
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.

Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

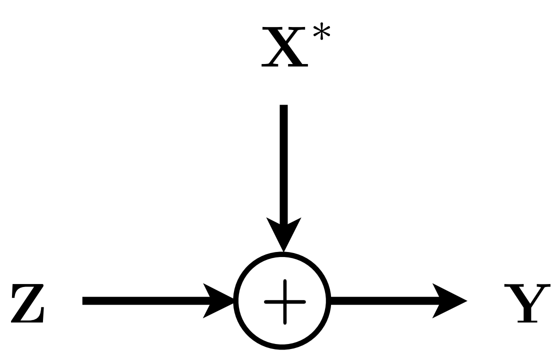
3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.



Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

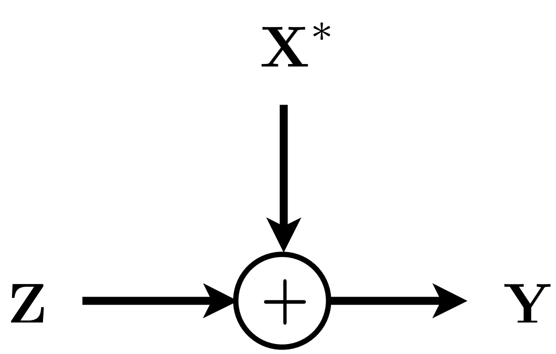
3. The theorem is proved by considering

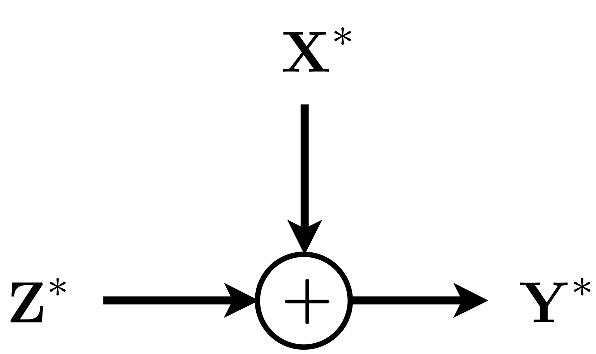
$$\begin{aligned}
&I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

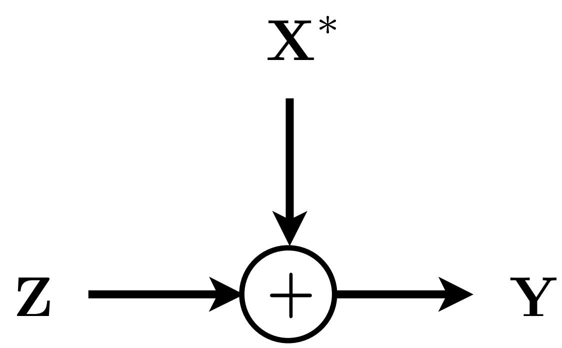
3. The theorem is proved by considering

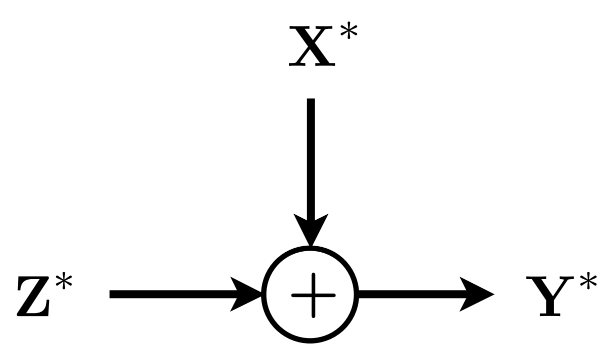
$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*})
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

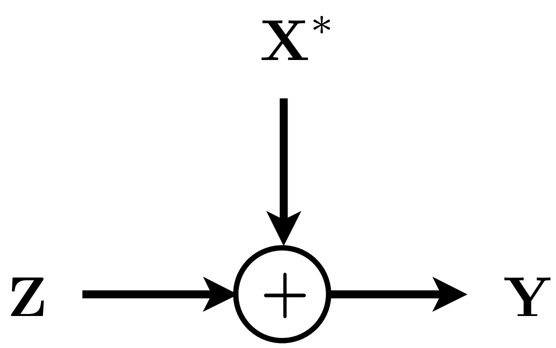
3. The theorem is proved by considering

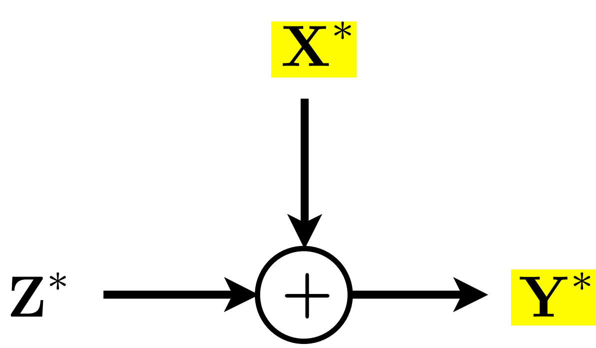
$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log\left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})}\right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log\left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})}\right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*}) \\
&\le 0.
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

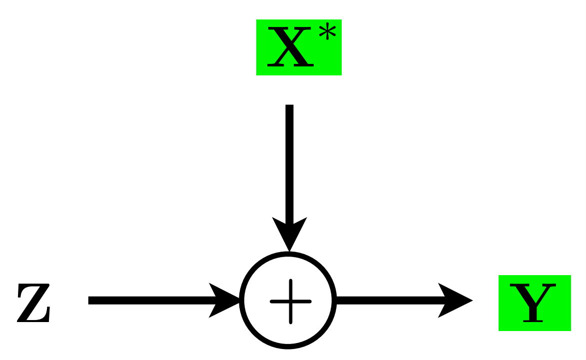
3. The theorem is proved by considering

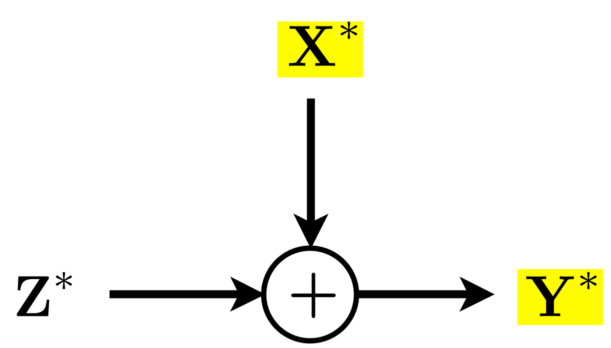
$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*}) \\
&\leq 0.
\end{aligned}$$

Theorem 11.32 For a fixed zero-mean Gaussian random vector $\mathbf{X}^*$, let

$$\mathbf{Y} = \mathbf{X}^* + \mathbf{Z},$$

where the joint pdf of $\mathbf{Z}$ exists and $\mathbf{Z}$ is independent of $\mathbf{X}^*$. Under the constraint that the correlation matrix of $\mathbf{Z}$ is equal to K, where K is any symmetric positive definite matrix, $I(\mathbf{X}^*;\mathbf{Y})$ is minimized if and only if $\mathbf{Z} = \mathbf{Z}^* \sim \mathcal{N}(0, K)$.





Proof

1. Since $E\mathbf{Z}^* = 0$, $\tilde{K}_{\mathbf{Z}^*} = K_{\mathbf{Z}^*} = K$. Therefore, $\mathbf{Z}^*$ and $\mathbf{Z}$ have the same correlation matrix.

2. By noting that $\mathbf{X}^*$ has zero mean, we apply Lemma 11.33 to see that $\mathbf{Y}^*$ and $\mathbf{Y}$ have the same correlation matrix.

3. The theorem is proved by considering

$$\begin{aligned}
& I(\mathbf{X}^*;\mathbf{Y}^*) - I(\mathbf{X}^*;\mathbf{Y}) \\
&= h(\mathbf{Y}^*) - h(\mathbf{Z}^*) - h(\mathbf{Y}) + h(\mathbf{Z}) \\
&= -\int f_{\mathbf{Y}^*}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int f_{\mathbf{Z}^*}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= -\int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}^*}(\mathbf{y}) d\mathbf{y} + \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}^*}(\mathbf{z}) d\mathbf{z} \\
&\quad + \int f_{\mathbf{Y}}(\mathbf{y}) \log f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} f_{\mathbf{Z}}(\mathbf{z}) \log f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= \int \log \left(\frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}^*}(\mathbf{y})} \right) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} - \int_{\mathcal{S}_{\mathbf{Z}}} \log \left(\frac{f_{\mathbf{Z}}(\mathbf{z})}{f_{\mathbf{Z}^*}(\mathbf{z})} \right) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \\
&= D(f_{\mathbf{Y}} \| f_{\mathbf{Y}^*}) - D(f_{\mathbf{Z}} \| f_{\mathbf{Z}^*}) \\
&\leq 0.
\end{aligned}$$