Consider the covariance matrix for a Gaussian with mean = (0, 0) and variance = $\sigma^2 \times I_2$ where σ^2 is a constant and I^2 is a 2 x 2 identity matrix.

a) Ans: What are the two principal components for this matrix and what are their eigenvalues?
b) Given a data point (x, y) from this distribution, what is the reconstructed data using the projection onto the first principal component of this matrix?
c) For this reconstructed value, what is the expected value of the reconstruction error?

a) Ans: To find the principal components and their corresponding eigenvalues for the given covariance matrix, we can follow these steps:

1. Write down the covariance matrix.
2. Calculate the eigenvalues and eigenvectors of the covariance matrix.
3. The eigenvectors represent the principal components, and the eigenvalues represent the variance along each principal component.

Given that the covariance matrix for a Gaussian distribution with mean $\mu = (0,0)$ and variance $\sigma^2 \times I_2$ is:

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

where I_2 is the 2x2 identity matrix.

To find the eigenvalues, we solve the characteristic equation

$\det(\Sigma - \lambda I) = 0$, where λ represents the eigenvalues:

$$\begin{Vmatrix} \sigma^2 - \lambda & 0 \\ 0 & \sigma^2 - \lambda \end{Vmatrix} = 0$$

$$(\sigma^2 - \lambda)^2 = 0$$

$$\lambda = \sigma^2$$

This tells us that both eigenvalues are σ^2, indicating that the variance along each principal component is σ^2.

To find the corresponding eigenvectors, we substitute each eigenvalue back into the equation

$(\Sigma - \lambda I)\mathbf{v} = \mathbf{0}$

and solve for $\mathbf{v}$. Since the covariance matrix is diagonal, the eigenvectors will be the standard basis vectors $(1,0)$ and $(0,1)$ for the given 2x2 covariance matrix.

So, the two principal components are the standard basis vectors $(1,0)$ and $(0,1)$, and their corresponding eigenvalues are both σ^2.

b) Ans: To reconstruct the data using the projection onto the first principal component of the covariance matrix, we need to project the data point onto the eigenvector corresponding to the largest eigenvalue. In this case, since both eigenvalues are σ^2, any eigenvector can be chosen as the first principal component.

Let's choose the eigenvector corresponding to the first principal component as

$\mathbf{v}_1 = (1,0)$.

Given a data point (x,y), the projection of this data point onto $\mathbf{v}1$ is given by the dot product of the data point and the eigenvector:

$$\text{Projected data} = (x,y)\cdot \mathbf{v}_1 = (x,y)\cdot(1,0) = x$$

So, the reconstructed data using the projection onto the first principal component is simply the x-coordinate of the original data point.

In summary, the reconstructed data point using the projection onto the first principal component is $(x,0)$.

c) Ans: The reconstruction error for a data point projected onto a principal component is the Euclidean distance between the original data point and its projection onto the principal component. Since we are projecting onto the first principal component, the reconstruction error for a data point (x,y) would be the difference between the original y value and the reconstructed y value (which is 0 in this case, since we're projecting onto the x-axis).

So, the reconstruction error for a data point (x,y) projected onto the first principal component would be:

$$\text{Reconstruction error} = |y-0| = |y|$$

Now, since the data is drawn from a Gaussian distribution with mean $\mu = (0,0)$ and variance $\sigma^2 \times I2$, the expected value of y (and thus the expected value of the reconstruction error) would be 0, because the mean of the distribution is 0.

Therefore, the expected value of the reconstruction error for a data point projected onto the first principal component is 0.