Image processing with Markov random fields
Introduction to Principal Component Analysis

Machine Learning II 2019

4.2.1 Medium-scale modelling of textures (1p)

Images of sizes 3×3 are not large enough to see the true texture generated by the simple homogenous Markov field. Adjust the model for 16×16 and 32×32 images and syntehsise some textures with different α, β, δ parameters. Note that we can set $\delta = 1$ as by scaling the pixel intensities by $\sqrt{c}$ we scale the entire inverse covariance matrix by the factor of c.

Solution

An example piece of code for finding the inverse covariance matrix (ignoring the scaling factor 1/2) for modelling the simple homogenous Markov random field.

```
def inverse_sigma(ndim, alpha, beta, delta):
    # Intensity deviation penalties
    V = np.diag([delta**2] * ndim**2)

    for row in range(ndim):
        for col in range(ndim):
            i = row*ndim + col

            # Horizontal penalties
            if col != ndim - 1:
                V[i, i + 1] -= alpha
                V[i + 1, i] -= alpha
                V[i, i] += alpha
                V[i + 1, i + 1] += alpha

            # Vertical penalties
            if row != ndim - 1:
                V[i, i + ndim] -= beta
                V[i + ndim, i] -= beta
                V[i, i] += beta
                V[i + ndim, i + ndim] += beta

    return V
```

5.1.1 Two-dimensional PCA (2p)

Implement the the PCA algorithm as a function for two-dimensional case. For that do the following steps. Centre the data and find the covariance matrix Σ. Use `scipy.linalg.eigh` function find eigenvectors of Σ. Apply the method on the distribution

$$\begin{cases} y_1 = 2x_1 + x_2 + 1 \\ y_2 = 2x_1 - x_2 - 3 \end{cases}$$

where $x_1 \sim \mathcal{N}(0,1)$ and $x_2 \sim \mathcal{N}(0,1)$. Complete the exercise by drawing a blue curve that represent empirical variance in each direction centered to the center of gravity of the data:

* In the tutorial above we draw the variance curve for centered data. Here you must draw the standard variance curve on the original data. As the probing lines for the variance go through the center of the gravity you need to shift the variance curve to the right place.

As you know the original data transformation you also know the theoretical variance curve. Draw this in red. Add also empirical center of gravity and the theoretical mean value. To see how much data is needed to adequate reconstruction of the distribution experiment with differen sample sizes. Interpret results.

Solution

Let $x = (x_1, x_2)^T$, then $x \sim \mathcal{N}(0, \mathrm{I})$. We can write $y = (y_1, y_2)^T$ in matrix form $y = Ax + \mu$ where $\mu = (1, -3)^T$ and $A = \begin{bmatrix} 2 & 1 \\ 2 & -1 \end{bmatrix}$. Then

$$y \sim \mathcal{N}(\mu, \Sigma = AA^T).$$

Consider now an arbitrary direction $w = (w_1, w_2)^T$. The theoretical variance of the data projected onto this direction is

$$\mathrm{Var}(w^T y) = \mathrm{Var}(w_1 y_1 + w_2 y_2) = w_1^2 \Sigma_{1,1} + w_2^2 \Sigma_{2,2} + 2 w_1 w_2 \Sigma_{1,2}.$$

Variance curves have been depicted in figure 1.

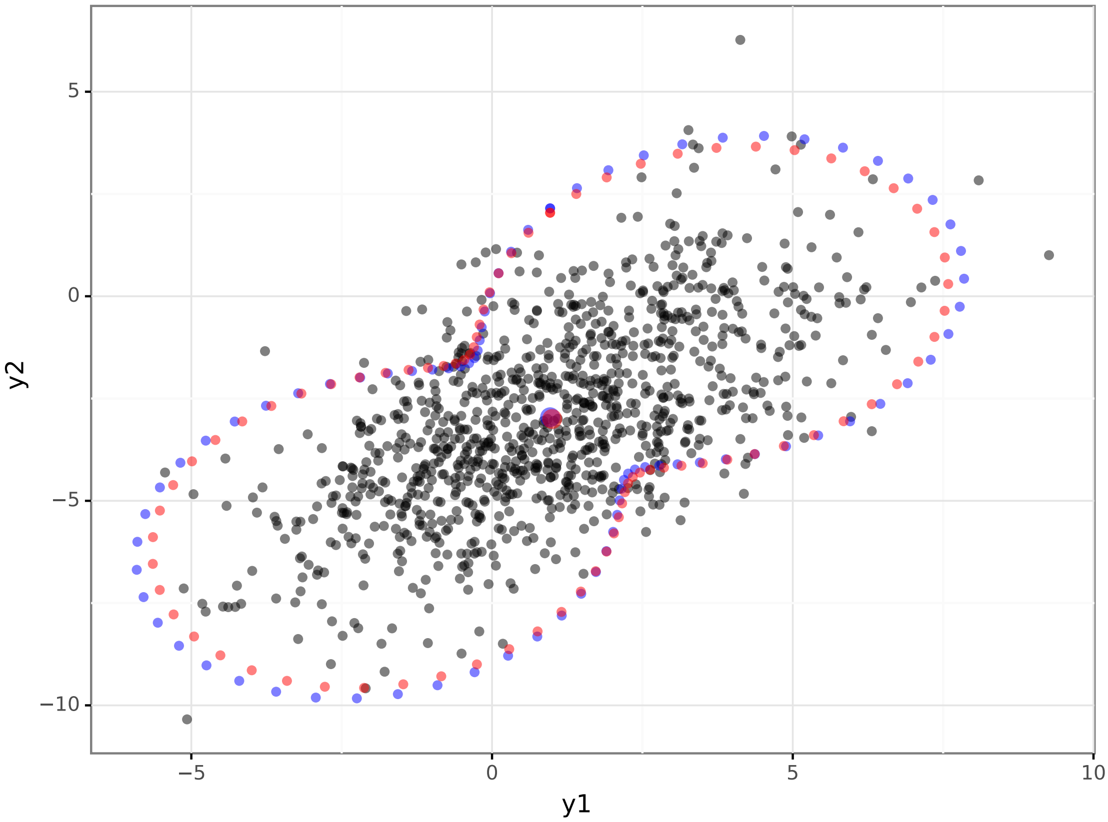

Figure 1: Variance curves with $n = 1000$. Empirical points in blue and theoretical points in red.