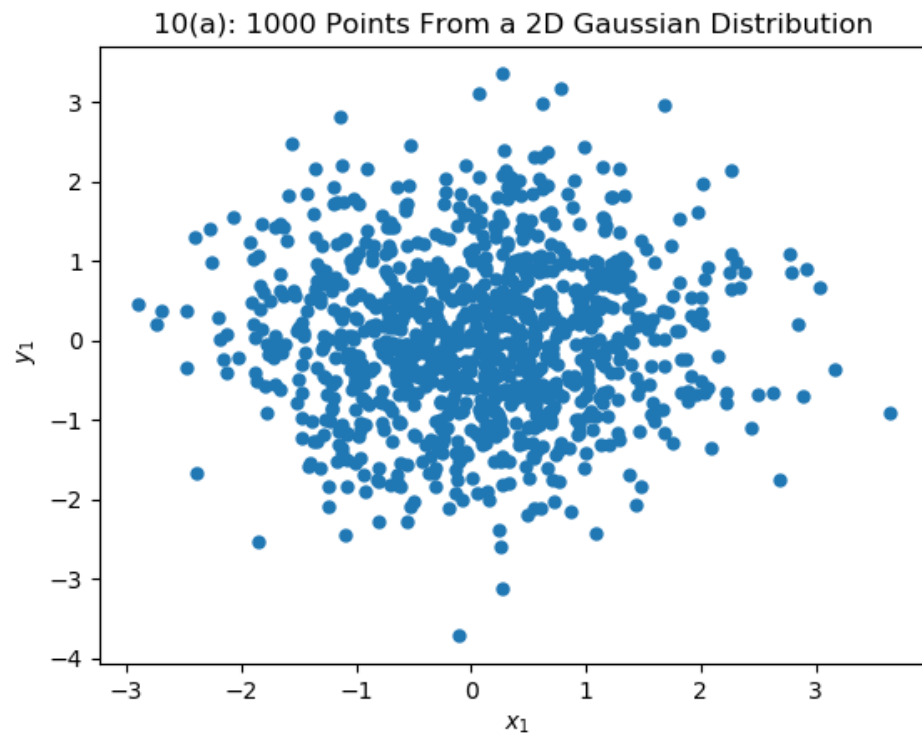
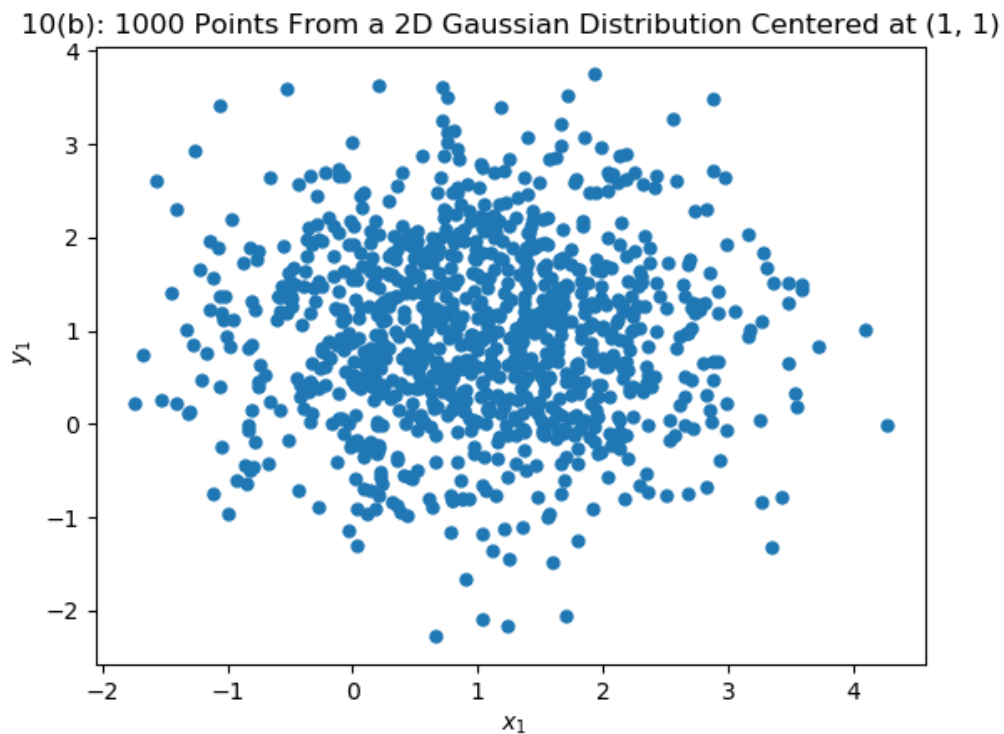


Problem Set 0 Coding Problems

10. (a)

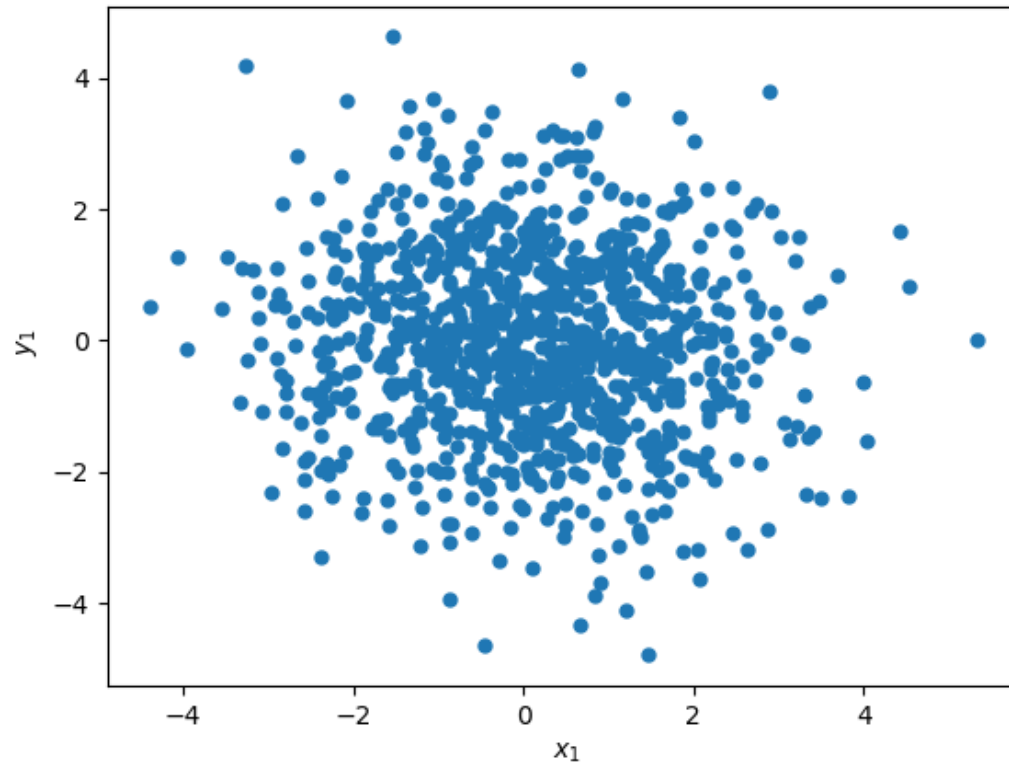


(b) The “center” of the set of points moves to roughly (1, 1).



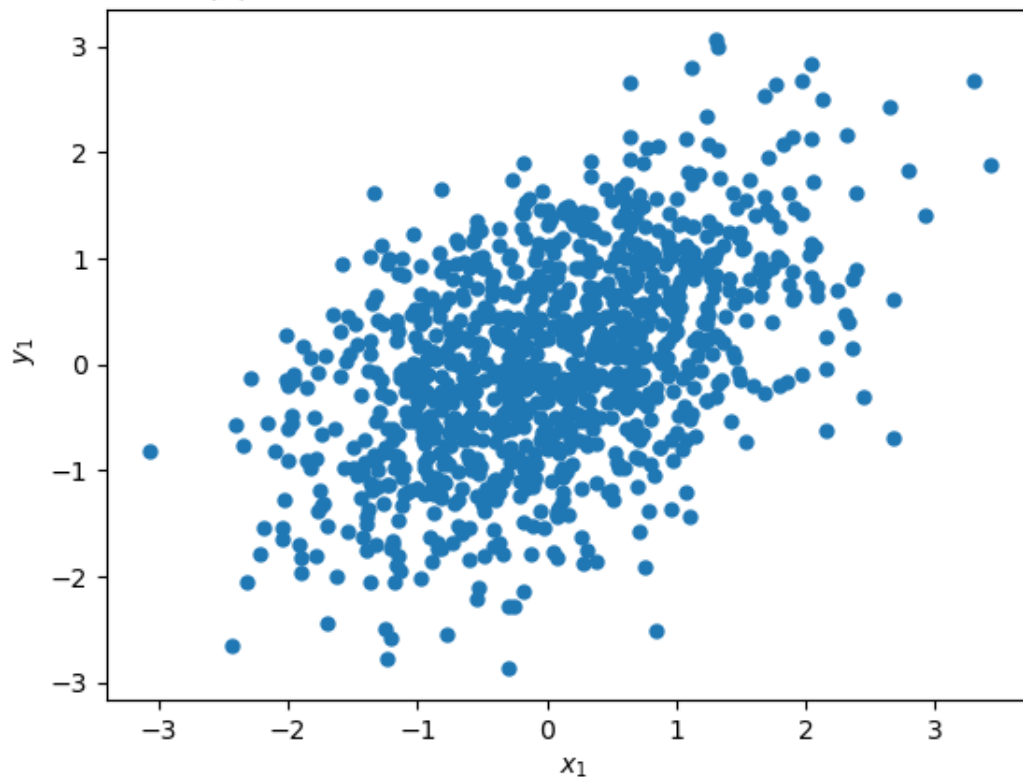
(c) The points become more spread apart.

10(c): 1000 Points From a 2D Gaussian Distribution

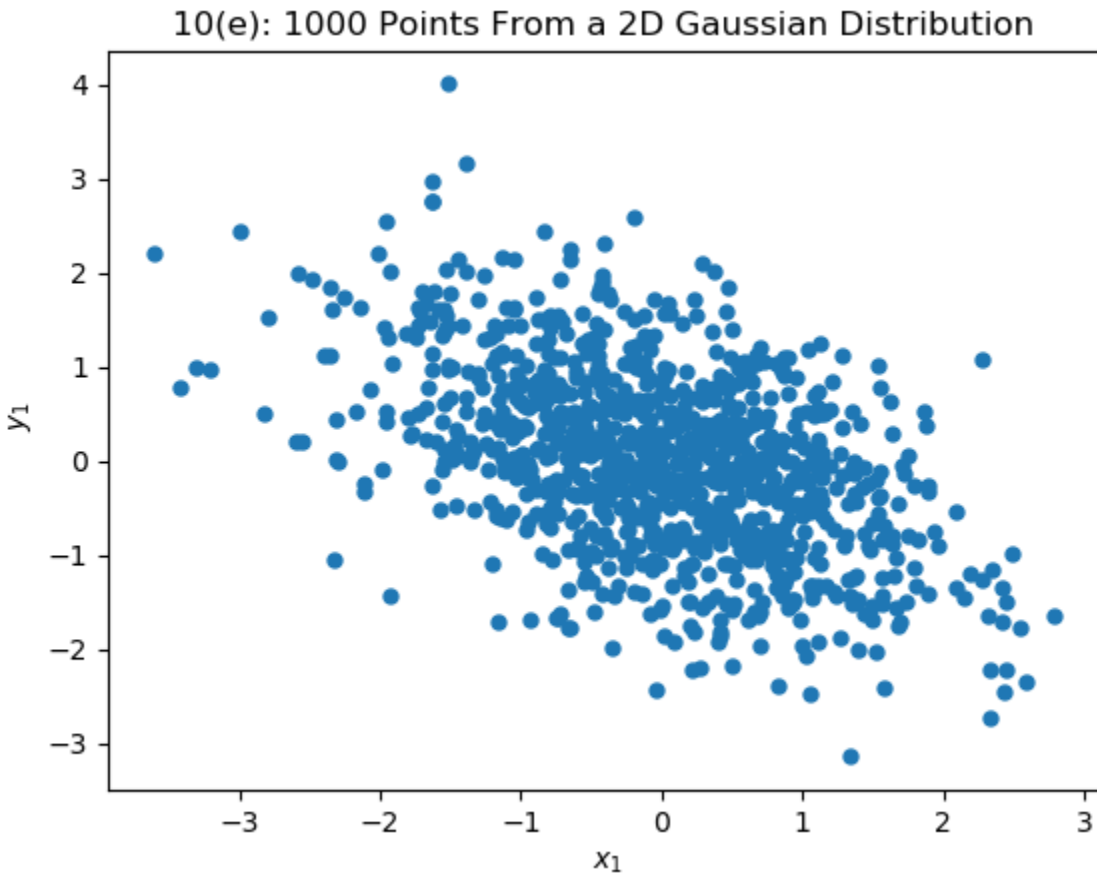


(d) The points now are positively correlated.

10(d): 1000 Points From a 2D Gaussian Distribution



(e) The points are now negatively correlated.



11. The eigenvector is $\langle 0, 1 \rangle$, corresponding to an eigenvalue of 3.

12.

(a) The name of the data set is “Large-scale CelebFaces Attributes (CelebA) Dataset,” provided by the Chinese University of Hong Kong.

(b) It can be obtained from <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>.

(c) The data set includes a large number of pictures of celebrities’ faces (10,177 unique identities). These images can be used for face recognition or detection.

(d) There are a total of 202,599 labelled images.

(e) There are a total of 40 features, including arched eyebrows, big nose, black hair, high cheekbones, male, female, no beard, narrow eyes, wavy hair, smiling, wearing hat, and so on.