Lecture 5

Exercises

Exercise 11. Download the `circles_32x32.mat` data set from D2L. It contains 1861 32×32 images, each with a circle drawn in it; see Figure 6 for a sampling of 16 of these images. All of the circles have the same radius. Therefore, the data set is two dimensional, since every image can be characterized completely by the center of the circle (which has two coordinates). Now run PCA on the data set, and examine the eigenvalues. What dimension do they indicate? What do you attribute this result to? Its high dimension, or the nonlinear way in which the 2D data was embedded into $\mathbb{R}^{1024}$ (or both, or something else)? Visualize the principal components. If you have taken harmonic analysis or are familiear with the Fourier transform, do they remind you of anything? (If not, do not worry about it. Just enjoy the cool pictures)

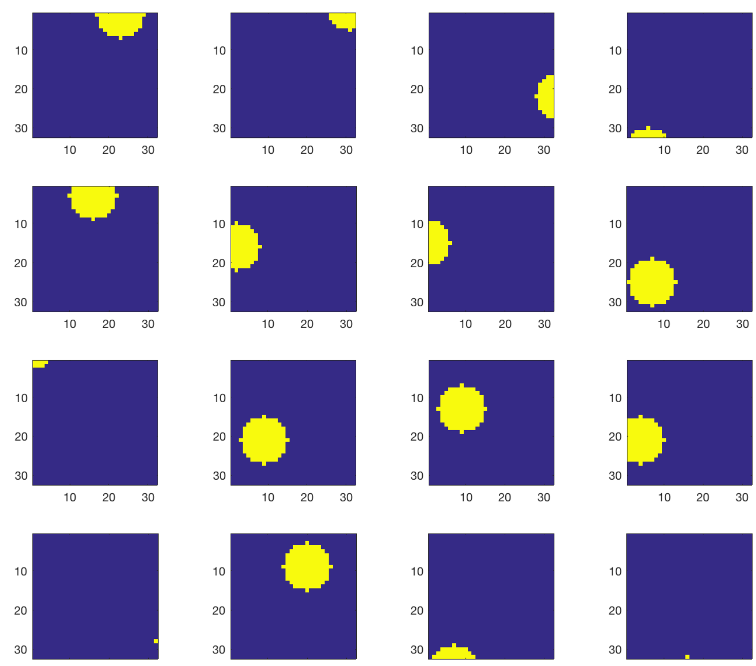

Figure 6: 16 images from the circles data set.

4 Primer on Hilbert Spaces

We use the notation:

$$\mathbb{N} = \{1, 2, 3, \dots\}.$$

4.1 Review of vector spaces, norms, inner products

A *real vector space* is a set V along with an addition "+" on V and a scalar multiplication "·" on V such that the following properties hold:

- Commutativity:
$$\forall\, u, v \in V, \quad u + v = v + u.$$
- Associativity:
$$\forall\, u, v, w \in V, \quad (u + v) + w = u + (v + w),$$
$$\forall\, v \in V, \forall\, a, b \in \mathbb{R}, \quad (ab)v = a(bv).$$
- Additive identity:
$$\text{there exists } 0 \in V \text{ such that } \forall\, v \in V, \quad v + 0 = 0 + v.$$
- Additive inverse:
$$\text{for each } v \in V, \text{ there exits } w \in V \text{ such that } v + w = 0.$$
- Multiplicative identity:
$$\forall\, v \in V, \quad 1v = v.$$
- Distributive properties:
$$\forall\, u, v \in V, \forall\, a \in \mathbb{R}, \quad a(u + v) = au + av,$$
$$\forall\, v \in V, \forall\, a, b \in \mathbb{R}, \quad (a + b)v = av + bv.$$

For our purposes, we will only discuss real vector spaces, and so henceforth shall just refer to them as "vector spaces." Every vector space has an *algebraic basis*. A set $\{v_1, \dots, v_n\} \subset V$ is *linearly independent* if

$$\forall\, a_1, \dots, a_n \in \mathbb{R}, \quad \sum_{i=1}^{n} a_i v_i = 0 \Longleftrightarrow a_1 = a_2 = \cdots = a_n = 0.$$

A set $\mathcal{B} = \{v_1, v_2, \dots\} \subset V$ is an *algebraic basis* for V if:

- Linear independence: Every finite subset $\{v_{k_1}, \ldots, v_{k_n}\} \subset \mathcal{B}$ is linearly independent.
- Spanning: For each $v \in V$, there is some $n \in \mathbb{N}$, $\{a_1, \ldots, a_n\} \subset \mathbb{R}$, and $\{v_{k_1}, \ldots, v_{k_n}\} \subset \mathcal{B}$ such that

$$v = \sum_{i=1}^{n} a_i v_{k_i}.$$

If $\mathcal{B}$ is finite, then the *dimension* of V is the number of vectors in $\mathcal{B}$. If $\mathcal{B}$ is infinite, then V is infinite dimensional.

A *normed space* is a vector space V with a norm $\|\cdot\| : V \to \mathbb{R}$, where the norm satisfies:

- Nonnegative:
$$\forall\, v \in V, \quad \|v\| \geq 0.$$
- Strictly positive:
$$\|v\| = 0 \iff v = 0.$$
- Homogeneous:
$$\forall\, v \in V, \forall\, a \in \mathbb{R}, \quad \|av\| = |a|\|v\|.$$
- Triangle inequality:
$$\forall\, u, v \in V, \quad \|u + v\| \leq \|u\| + \|v\|.$$

An *inner product space* is a vector space V with an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$, where the inner product satisfies:

- Nonnegative:
$$\forall\, v \in V, \quad \langle v, v \rangle \geq 0.$$
- Strictly positive:
$$\langle v, v \rangle = 0 \iff v = 0.$$
- Additivity:
$$\forall\, u, v, w \in V, \quad \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle.$$

- Homogeneous:

$$\forall u, v \in V, \forall a \in \mathbb{R}, \quad \langle au, v \rangle = a\langle u, v \rangle.$$

- Symmetry:

$$\forall u, v \in V, \quad \langle u, v \rangle = \langle v, u \rangle.$$

Every inner product space defines a normed space by taking $\|v\| = \sqrt{\langle v, v \rangle}$. Even in infinite dimensional inner product spaces, the *Cauchy-Schwarz inequality* still holds:

$$\forall u, v \in V, \quad |\langle u, v \rangle| \leq \|u\| \|v\|.$$

4.2 Hilbert spaces

Adapted from [1, Appendix B.3]

A sequence $(v_i)_{i \in \mathbb{N}} = (v_1, v_2, \ldots)$ in a normed space $(V, \|\cdot\|)$ is a *Cauchy sequence* if for every $\epsilon > 0$, there exists an $n \in \mathbb{N}$ such that for all $i, j > n$, $\|v_i - v_j\| < \epsilon$.

A sequence $(v_i)_{i \in \mathbb{N}}$ *converges* to a point $v \in V$ if for every $\epsilon > 0$, there exists an $n \in \mathbb{N}$ such that $\|v_i - v\| < \epsilon$. We write $v_i \to v$ as $i \to \infty$.

A normed space $(V, \|\cdot\|)$ is *complete* if each Cauchy sequence in the space converges to some point in V.

A *Banach space* is a complete normed space.

A *Hilbert space* is a complete inner product space. Hilbert spaces are often denoted $\mathcal{H}$.

References

[1] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Adaptive Computation and Machine Learning. The MIT Press, 2002.

[2] Afonso S. Bandeira. Ten lectures and forty-two open problems in the mathematics of data science. MIT course *Topics in Mathematics of Data Science*, 2015.

[3] Jon Shlens. A tutorial on principal component analysis. arXiv:1404.1100, 2014.

[4] Karl Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, Series 6*, 2(11):559–572, 1901.

[5] V. A. Marchenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72(114):507–536, 1967.

[6] J. Baik, G. Ben-Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005.

[7] Debashis Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, pages 1617–1642, 2007.