PCA algorithms proof

Friday, June 26, 2020 1:00 AM

Suppose we have the following collection of m points

$\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$ in $\mathbb{R}^n$

We like to reduce each $x^{(i)}$ in the collection, i.e,

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^l \qquad l \leq n$$

$$f(x) = c$$

So f is a function that map x to a less dimension vector space, so f is our coding function. That means we'll need a decode function for f g

$$g : \mathbb{R}^l \longrightarrow \mathbb{R}^n \qquad l \leq n \quad \text{and} \quad g(f(x)) \approx x$$

the most simple transformation is define $g(c) = Dc$, $D \in$

to keep the decoding function easy let's define some properties for D

1) the columns of D define an orthonormal basis
let be $d^{(1)}, d^{(2)}, \dots, d^{(l)}$

$$d^{(i)T} d^{(j)} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

2) $D^T D = I_l$

this properties give the problem a unique solution

Now we need a definition for $f(x) = c$, and we know

$g(c) = Dc$ and $g(f(x)) \approx x$ ...

problem

$$c^* = \underset{c}{\arg\min} \; \| x - g(c) \|_2$$

because L2 norm is alway positive then optimize the
before is the same that optimize the square of
because square function are monotonic for positive arg

$$c^* = \underset{c}{\arg\min} \; \| x - g(c) \|_2^2$$

$$\| x - g(c) \|_2^2 = (x - g(c))^T . (x - g(c))$$
$$= x^T x - x^T g(c) - g(c)^T x + g(c)^T g(c)$$

$x^T g(c) \in \mathbb{R}$

$\Rightarrow x^T g(c) = (x^T g(c))^T \Rightarrow = x^T x - 2x^T g(c) + g$

$= g(c)^T x \Longrightarrow x^T x$ doesn't depend so is constant

$\Rightarrow c^* = \underset{c}{\arg\min} \left(-2x^T g(c) + \right.$

Using the definition of $g(c) = Dc$

$\Rightarrow c^* = \underset{c}{\arg\min} \left(-2x^T Dc + (Dc)^T \right.$