Presentation on Variational Characterization

Tianlin Liu

Feb 12 2015

Theorem (Variational Characterisation of SVD)**.** *Let $A = U\Sigma V^T$ be a Singular Value Decomposition(SVD). Let $u \in R^n$ and $v \in R^k$ be unit vectors. Then the maximum value of the function $f(u, v) = u^T Av$ over the set of u and v is the largest singular value of A.*

1 SVD Review

Let's have a quick review on SVD!

Definition.

$A = U\Sigma V^T$ is called the **SVD** of A if U and V are orthogonal matrix, and Σ is a diagonal matrix with non-negative entries. The vectors $v_1, \ldots, v_k$ are the **right singular vectors**, and the vectors $u_1, \ldots, u_n$ are the **left singular vectors**.

Definition.

The diagonal entries of Σ, denoted $\sigma_1, \ldots, \sigma_k$, are the **singular values** of SVD of A.

2 Lagrange Multipliers Principle

Theorem.

f has a local minimum or maximum on the set $v \in V : g(v) = 0, g \in C^1(\mathcal{V}, R)$ only if there exists $\lambda \in \mathbb{R}$ such that

$$d(f) = \lambda dg(v)$$

Remark. *We will not give a proof for this theorem. But in class I will talk about the motivation and intuition behind it.*

3 Proof of Variational Characterisation of SVD

Before really start proving, let's talk about the central ideas. So far, to prove this theorem, we have nothing but 3 main tools:

1. The notion of **optimization**(the idea introduced by CoCo).
2. The notion of **SVD form**(the idea that introduced by Denis).
3. The notion of **Lagrange Multipliers Principle**(the idea I just introduced).

Proof. The proof is simply the usage of these 3 tools!

Firstly, we give f two scalar constraints as suggested by Lagrange Multipliers Principle. Then we have a optimization problem:

$$g_1(u, v) \equiv u^T u - 1 = 0 \quad \text{and} \quad g_2(u, v) \equiv v^T v - 1 = 0$$

Further, by vector calculus,

$$df(u, v)[\delta u, \delta v] = v^T A^T \delta u + u^T A \delta v$$

$$dg_1(u, v)[\delta u, \delta v] = 2u^T \delta u$$
$$dg_2(u, v)[\delta u, \delta v] = 2v^T \delta v$$

By the Lagrange Multiplier Principle, after substituting and rearranging terms, we have

$$v^T A^T \delta u - 2\lambda_1 u^T \delta u) + (u^T A \delta v - 2\lambda_2 v^T \delta v = 0$$

Note that δu and δv can be chosen independently,

$$v^T A^T = 2\lambda_1 u^T \quad \text{and} \quad u^T A = 2\lambda_2 v^T$$

Right-multiplying the first equality with u and the second equality with v, we conclude that

$$u^T A v = 2\lambda_1 = 2\lambda_2 \equiv \sigma$$

as desired.

□

Remark.

Recall that from Analysis class, we know that a real valued function on a compact subset of R^n has a maximum and minimum. Hence the existence of a solution is clear. However, the solution may not be unique.

If time permits I will talk about the idea of **dyadic form of the SVD**, in which we shall write $A = \sum_{j=1}^{n} \sigma_j u_j v_j$. With this concept, it is easy to verify the two claims in page 13 of the scripts to be true.

4 References

[1] Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[2]http://www.math.vt.edu/people/embree/cmda3606/chapter8.pdf

[3]http://www.math.uconn.edu/∼leykekhman/courses/MATH3795/Lectures/Lecture _9_Linear_least_squares_SVD.pdf

[4] http://www.caam.rice.edu/∼caam440/