Rayleigh-Ritz Teoremi

Sentetik görüntü algoritmasını gösterdiğimizde, Rayleigh-Ritz kuramına atıf yapmıştık. Bu yazıda bütün kuramın ispatını veriyoruz. İspatta kullanılan küme sanal sayılar kümesidir. Bizim örneğimiz için gerçek sayılar kümesi kullanılıyor, fakat aynı ispat hala geçerli olacak.

Problem

Bir kare matrisin özdeğerlerini büyüklük sırasına dizersek, bu değerlerin kısıtlı bir minimizasyon / maksimizasyon probleminin çözümü olduğun görüyoruz. Kısıtlı derken, $x * x$ (x vektör devriği çarpı x, yani x'in uzunluğu) çarpımını 1'e kısıtlı tutmaktan bahsediyorum. Böylece maksimizasyon problemimizin sonsuzluğa gitmesini engellemiş oluyoruz. λ sembolu genelde özdeğerler için kullanılır. Yıldız işareti * ise sanal sayılar uzayında, devrik yapmak demektir. Gerçek sayılar uzayında olsaydık, o zaman T işaretini kullanabilirdik. (T transpose kelimesinden gelir).

$\forall x \in C^n$

$\lambda_1 x^* x \leqslant x^* A x \leqslant \lambda_n x^* x$

$$\lambda_{ust} = \lambda_n = \max_{x^*x=1} (\frac{x^* A x}{x^* x}) = \max_{x^*x=1} (x^* A x)$$

$$\lambda_{alt} = \lambda_n = \max_{x^*x=1} (\frac{x^* A x}{x^* x}) = \max_{x^*x=1} (x^* A x)$$

Problemi üstte tanımladıktan sonra, ispatına gelelim.

A matrisi, Hermit matrisi olduğu için, elimizde bu A matrisine tekabül eden birincil (unitary) bir matris var demektir. Bu birincil matrisi U ile temsil edersek, şu sonuca da varırız.

$A = U \Lambda U^*$

$\Lambda = diag(\lambda_1 \lambda_2 ..., \lambda_n)$

Bu demektir ki,

$\forall x \in C^n$

$x^* A x = x^* U \Lambda U^* x = (U^* x)^* \Lambda (U^* x)$

$$\sum_{i=1}^{n} \lambda_i |(U^* x)_i|^2$$

Ufak iki not olarak düşmek gerekiyor. Yukarıdaki 3. eşitliğe gelmemizin sebebi aşağıdakinin doğru olmasıdır.

$x^* U = (U^* x)^*$

Doğrusal cebirde bilinen çevirimlerden biridir bu. En son not olarak, toplamlı eşitliğe gelebilmemizin sebebi (4. terim) şundandır. $U^* x$ yerine W koyarsak,

W^*W çarpımının her zaman W'nin uzunluğunu verir. Yani bir vektörün uzunluğunu bulmak için vektörün devriğini kendisi ile çarpmak gerekir, bu çarpım uzunluğun karesidir.

Devam ediyoruz. Her $|(U^*x)_i|^2$ ifadesi artı değerli olmaya mecbur olduğu için,

$$\lambda_{alt} \sum_{i=1}^{n} |(U^*x)_i|^2 \leqslant x^*Ax = \sum_{i=1} \lambda_i |(U^*x)|^2 \leqslant \lambda_{ust} \sum_{i=1}^{n} |(U^*x)_i|^2 \leqslant x^*Ax$$

Üstteki eşitsizliğin doğru olmasının bir sebebi var. Elimizde 3 tane değişik 1..n arası yapılan toplam var. Dikkatle bakarsanız, ortadaki toplam içinde i ile kontrol edilen, bütün özdeğerlerin toplandığını göreceksiniz. Buna kıyasla mesela en soldaki, toplam içinde sürekli aynı 'alt özdeğer' toplandığını farketmemiz lazım. Buna bakarak anlıyoruz ki, tabii ki bütün özdeğerlerin toplamı, tekrar eden aynı özdeğer değerinin toplamından fazla olacaktır! Çünkü iki tarafta da özdeğerler haricindeki bütün terimler birbirine eşit. Daha da basitleştirmek için U'yu yokedelim.

U'da birincil bir matris olduğu için,

$$\sum_{i=1}^{n} |(U^*x)_i|^2 \sum_{i=1} |x_i|^2 = X^*x$$

cunku

$|U^*x| = |x|$

Ispat

$|U^*x| = (U^*x)^*(U^*x) = x^*UU^*x = x^*x = |x|$

Böylece göstermiş oluyoruz ki,

$\lambda_1 x^*x \leqslant \lambda_{alt} x^*x \leqslant x^*Ax \leqslant \lambda_{ust} x^*x$