특이값 분해2

특이값 분해를 계산하는 방법

- eigendecomposition을 사용하여 SVD를 계산한다.
- A가 5x3 행렬이라고 해보자.
    - AA^T: 5x5, A^TA: 3x3

$$AA^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T = U\Sigma^2 U^T$$
$$A^TA = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma U^T = V\Sigma^2 V^T$$

    - V는 orthonormal한 컬럼을 가지는 행렬이다. 따라서, V^TV는 사라진다.
    - D(=Σ)가 Diagonal matrix 일 때, ⇒ $DD^T = D^2$
    - U, V가 orthonormal한 컬럼들을 가지는 정사각행렬이라고 할 때,
      $U^T = U^{-1}$ 이므로, eigendecomposition과 유사한 형태로 볼 수 있다.
    - SVD를 풀기 위해, AA^T를 만들고, 이를 eigendecomposition으로 만들어서 풀게된다.

SVD가 만족해야하는 조건

1. U, V 행렬들이 orthonormal eigenvector로 이루어짐
2. Σ^2의 eigenvalue가 양수

   ⇒ 음수의 루트를 수행하면 실수 범위를 벗어나므로
3. AA^T, A^TA로부터 얻을 수 있는 Σ^2가 서로 동일해야한다.

+ AA^T, A^TA가 diagonalizable

조건1) U, V 행렬들이 orthonormal eigenvector로 이루어짐

Diagonalization of Symmetric Matrices

- 용어
    - **Diagonalizable**: NxN 행렬이 있을 때, N개의 선형독립인 eigenvector가 존재해야한다.
    - **Symmetric**: diagonal line을 기준으로 transpose 시켰을 때, 동일한 행렬이 나온다.
- eigendecomposition/diagonalizable하기 위한 조건: “symmetric”
    - $A = A^T$
    - ⇒ $(AA^T)^T = A^TA$
- 정사각행렬이고 Symmetric 조건을 만족하면, 이 행렬은 항상 orthogonally diagonalizable한다.

Spectral Theorem of Symmetric Matrices

→ symmetric matrix의 특성: Spectral Theorem

1. **eigenvalues는 반드시 실근이 나와야한다.(중근 허용)**
    - symetric matrix A에 대해, $det(A - \lambda I) = 0$ 에서 절대 허근이 나오지 않는다.
    - 허근이 나오면, N개의 선형독립인 eigenvector를 뽑을 수 없게 된다.
      허근에 해당하는 λ가 나와도, 허근들은 고려하지 않기 때문이다.
2. 각각의 실근에 해당하는 eigenvalue들마다 나온 중근의 개수에 해당하는 최대의 개수만큼 eigenspace의 dimension/basis가 뽑혀져 나와야 한다.

- algebraic multiplicity = geometric multiplicity 이어야 한다.
- 용어
  - algebraic multiplicity: 방정식의 각 근의 중근의 개수
  - geometric multiplicity: eigenspace의 basis 개수
  - multiplicity: 중근성, 중근의 수

3. **eigenspace가 수직이다.**(= eigenvector들이 수직이다. 이 space를 이루는 벡터들이 서로 수직이다.)

   = eigenspace 내에서 orthonormal한 벡터를 뽑을 수 있다.

   - eigenvalue가 다른 eigenspace는 그 basis가 전부 직교하지 않는 경우가 존재한다.
     그러나, symmetric matrix의 서로 다른 eigenvector들이 모두 수직을 만족한다.
     - 여기서 "eigenvector" 대신, "eigenvalue가 다른 eigenspace"라고 부른 이유는, 하나의 eigenvalue로 나타날 수 있는 eigenspace가 2차원 이상일 수 있기 때문이다.

Spectral Decomposition

Q. 정사각행렬이며, symmetric matrix인 경우, eigendecomposition은 어떻게 나올까?

- 수식

$$\begin{aligned} \bullet\ S = UDU^{-1} = UDU^T &= [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= [\lambda_1\mathbf{u}_1 \quad \lambda_2\mathbf{u}_2 \quad \cdots \quad \lambda_n\mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T \end{aligned}$$

▼ 설명(그림)

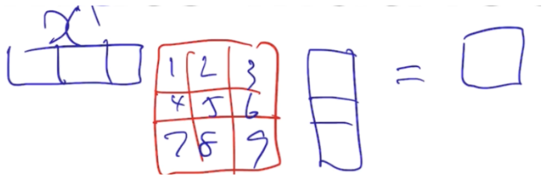
$$\begin{aligned} UDU^T &= [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= [\lambda_1\mathbf{u}_1 \quad \lambda_2\mathbf{u}_2 \quad \cdots \quad \lambda_n\mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T \end{aligned}$$

- λ: eigenvalue
- S는 symmetric matrix이므로, U는 orthonormal eigenvector N개 ⇒ so, $U^T = U^{-1}$
- λ_1은 u_1과 u_1^T의 outer product에만 영향을 준다.
- U는 orthonormal vectors로 이루어져 있으므로, u_1과 u_1^T의 outer product의 크기는 1이다.
- $\lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots$
  ⇒ 길이가 1인 벡터들을 방향만 달리하여 만든 것에 각각 λ를 곱하여 scaling한 것으로 해석할 수 있다.
- 결론) 주어진 symmetric matrix는 rank1의 outer product로 이루어져 있다.
  (길이 1인 벡터를 곱하고, 늘리고, 더하는 과정으로 복원 가능)

조건2) Σ^2 의 eigenvalue가 양수이다.

Positive Definite Matrices

- 지금까지 알아낸 것
    - Symmetric matrix에서 eigenvectors가 N개의 선형독립이 나오고, 그들이 모두 otrhonormal한 eigenvector들이 나온다.
    - U와 V는 orthonormal한 eigenvectors가 뽑히므로, 우리가 구한 eigendecomposition에서의 U, V의 역할이 가능함을 보았음
- 해결해야하는 것?
    - AA^T, A^TA의 eigenvalues가 모두 0보다 커야한다는 조건
    - ⇒ Positive definite, Positive semi-definite(0과 같은 경우도 포함) 라는 특성임
- 정사각행렬이 주어졌을 때, 아래의 경우, 상수 output이 나온다.



    - 어떤 특수한 matrix의 경우, 스칼라 값이 항상 0보다 큰 결과를 내게 되면, 해당 행렬을 positive definite라고 부른다.
    - 2차함수의 경우, x에 어떤 값을 넣어도 양수가 나오게 되는 경우가 있다. positive semi-definite의 경우, 축과 접해있는 경우를 생각할 수 있다.
    - Positive semi-definite인 경우, eigendecomposition, eigenvalue, eigenvector 측면에서 eigenvalue가 모두 양수가 나온다는 특징이 있다. (증명생략)
- Positive definite: $x^TAx > 0$
- Positive semi-definite: $x^TAx \geq 0$

정사각행렬인 경우, Positive definite 조건을 만족하는 경우?

Positive definite인 경우, diagonalization 혹은 eigendecomposition이 가능 한지 여부에 대해, 그럴수도있고, 아닐수도 있다. 그러나 찾아낸 eigenvalues는 모두 양수임

N개를 모두 찾을 수 없더라고, 찾아낸 것은 모두 양수라는 것

symmetric하다는 것 때문에 eigendecomposition이 언제나 존재하고 그것의 식은 아래와 같다.

$$\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

질량이 1짜리 성분에를 람다를 곱해서 더해주는 것과 같다.

가중치와 같이 해석이 가능하다.

그러나 람다에 음수가 들어가면, 질량으로 해석하기가 어렵게 된다.

따라서, symmetric + positive definite 조건이 합쳐지면, 양적으로 해석하기 용이해진다.

- 다시 SVD 계산으로 돌아와보자.

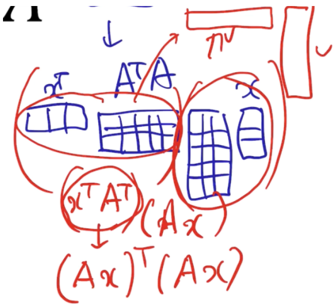
$$AA^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma\Sigma^T U^T = U\Sigma^2 U^T$$
$$A^TA = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T \Sigma U^T = V\Sigma^2 V^T$$

위의 두 AA^T, A^TA는 아래의 두 조건을 만족한다.

- Symmetric: $(AA^T)^T = AA^T$ and $(A^TA)^T = A^TA$
- Positive-(semi-)definite
  - $\mathbf{x}^T AA^T \mathbf{x} = (A^T\mathbf{x})^T(A^T\mathbf{x}) = \|A^T\mathbf{x}\|^2 \geq 0$
  - $\mathbf{x}^T A^TA \mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$

- 위의 식이 positive-(semi-)definite함을 증명
  - $x^T(AA^T)x$ 는 결국 자기 자신을 내적한 것과 같다.

    = norm의 제곱과 같으므로, 언제나 0보다 크거나 같다.



조건3) AA^T, A^TA로부터 얻을 수 있는 Σ^2가 서로 동일해야한다. (생략)

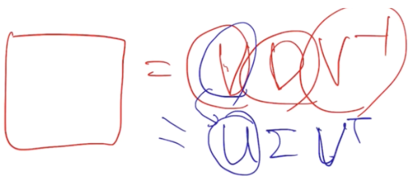
결론

SVD는 주어직 직사각행렬을 AA^T, A^TA를 각각 eigendecomposition해서 나온 어떤 eigenvector들을 U와 V로 쓰고, 두개의 공통적으로 나온 eigenvalue들을 항상 똑같은 세트가 나오게 된다.

그래서 이들을 시그마 스퀘어로 써서 그것들의 각각을 루트만 취하면, 원래의 U시그마 V^T라는 오리지널 A라는 직사각행렬의 SVD을 구할 수 있다.

눈여겨볼만한 특성

- 어떤 직사각행렬이 주어지든, SVD는 항상 존재
- eigendecomposition은 정사각행렬에서 존재할때, 안할때(역행렬 없는 때)가 있었음, SVD는 언제나 존재



- eigendecomposition은 동일한 행렬의 역행렬을 써야하므로 자유도가 SVD에 비해 떨어짐

- 이미 정사각행렬이고, 이미 symm이고, 이미 posit이면,
  이때의 eigendecomposition은 언제나 존재하고, 이에 역행렬 관계
- N개의 수직인 eigenvectors 가짐
- $U = V$로 동일하게 처리 가능, $V^T = V^{-1}$이므로, eigen-decomposition가 SVD에서 만족하는 모든 조건을 만족하게 됨 따라서 동일하게 볼 수 있다. 이 경우, eigen-decomposition가 SVD 역할 수행 가능