Note 7

07 Decompositions

by Alvin Wan

In this section, we consider two matrix decompositions. Before delving into either, we will need some terminology. We note the following definitions:

Matrix A is **positive semi-definite (PSD)** if the following three equivalent conditions hold:

(1) quadratic form is non-negative, $\forall x, x^T Ax \geq 0$

(2) all eigenvalues of A are non-negative, $\forall i, \lambda_i(A) \geq 0$

(3) a **Cholesky decomposition** exists, $\exists B, A = B^T B$

Note that PSD also implies eigenvalues and singular values of A are identical. We will prove this last fact in section 3.

Proof: We need to show that all three conditions are equivalent. To achieve this, we

(1) $\implies$ (2): Take $x = v$ of A. Then, $v^T Av = v^T \lambda v = \lambda \|v\|_2^2 \geq 0$ iff $\lambda \geq 0$.

(2) $\implies$ (3): Take the eigen decomposition of A, which is discussed below: $A = PDP^T = PD^{1/2}D^{1/2}P^T = (D^{1/2}P^T)^T(D^{1/2}P^T) = B^T B$, where $B = D^{1/2}P^T$.

(3) $\implies$ (1): $\forall x, x^T Ax = x^T B^T Bx = (Bx)^T Bx = \|Bx\|_2^2 \geq 0$

Matrix A is **positive definite (PD)** if $\forall x, x^T Ax > 0$; this is iff all eigenvalues of A are positive. Proof of this statement is nearly identical to the proof above.

1 Eigenvalue decomposition (EVD)

This decomposition may be more familiar as "diagonalization" or "spectral decomposition". In any case, we decompose a matrix A into its eigenvectors v_i and corresponding eigenvalues λ_i such that $Av_i = \lambda v_i$.

1.1 Intuition Review

This formulation tells us that A, as an operator, scales v_i by some scalar quantity, positive or negative - in effect extending the vector or flipping it over the origin.

The **spectral theorem** states that every real, symmetric $n \times n$ matrix has n eigenvectors s.t. $\forall i \neq j, v_i^T v_j = 0$. (See proof.) This has a few implications:

1. The characteristic polynomial of any matrix and thus its eigenvalues - including their multiplicities - are unique for each matrix.
2. There may be more than n eigenvectors; we choose a set of orthogonal vectors that span the eigenbasis.

Equipped with the spectral theorem, we can now construct the decomposition. Consider eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ with corresponding eigenvectors $v_1, v_2 \ldots v_n$ for matrix A, where $D = \text{diag}(\lambda_1, \lambda_2 \ldots, \lambda_n)$ and v_i form the column vectors of P. We then have that $A = PDP^{-1}$. Note that if all v_i are orthonormal, or where $\|v_i\|_2^2 = 1$ and $\forall i \neq j, v_i^T v_j = 0$, then $P^T P = I$, implying $P^T = P^{-1}$. Thus for orthonormal v_i,

$$A = PDP^T$$

Theorem If λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k.

Proof: Observe $A^2 v_i = AAv_i = A(\lambda_i v_i) = \lambda_i A v_i = \lambda_i^2 v_i$. Apply inductively to obtain result.

Applying the spectral theorem, we see the following. Note $P^T P = I$.

$$A^k = (PDP^T)^k = PDP^T PDP^T \cdots PDP^T = PD^k P$$

With this, we know for $A = M^{1/2}$, we have $A^2 = PD^2 P^T = M$, so $M = PD^{1/2} P^T$.

Theorem If A is invertible, then eigenvector v_i with eigenvalue λ_i of A is also an eigenvector of A^{-1} with eigenvalue λ_i^{-1}.

Proof: $Av_i = \lambda_i v_i \implies v_i = \lambda_i A^{-1} v_i \implies \lambda_i^{-1} v_i = A^{-1} v_i$

With the spectral theorem, $A^{-1} = (PDP^T)^{-1} = PD^{-1} P^T$.

1.2 Ellipsoids

First, consider the set all vectors that have length 1. This set gives rise to the unit ball.

$$\{x : \|x\|_2^2 \leq 1\}$$

Now, take the quadratic form $x^T Ax$, where A is positive semidefinite (PSD). We will show that the set of all vectors such that $x^T Ax$ has length 1 gives rise to a distortion of the unit ball, or the **ellipsoid**.

$$\{x : x^T Ax \leq 1\}$$

We can take the cholesky decomposition of $A = B^T B$, which was proved to exist in the exposition for this note. We then have the following.

$$x^T Ax = x^T B^T Bx = (Bx)^T (Bx) = \|Bx\|_2^2$$

.

Consider an eigenvector, eigenvalue pair v_B, λ_B of B. Say $x = v_B$. We then show that the quadratic form $x^T Ax$ has length $\frac{1}{\sqrt{\lambda_B}}$ in the direction of $x = v_B$.

$$x^T Ax = \|Bx\|_2^2 = \|Bv_B\|_2^2 = \|\lambda_B v_B\|_2^2 = \lambda_B^2 \|v_B\|_2^2$$

To find the boundary of our shape, we take $x^T Ax = \lambda_B^2 \|v_B\|_2^2 = 1$ instead of an inequality. The last part of the line above tells us that

$$\|v_B\|_2 = \frac{1}{\lambda_B}$$

We need to relate the eigenvalue of B, λ_B, back to the eigenvalue of A. So, take the eigen decomposition of $B = P_B D_B P_B^T$ and $A = P_A D_A P_A^T$.

$$A = B^T B = P_B D_B P_B^T P_B D_B P_B^T = P_B D_B^2 P_B^T = P_A D_A P_A^T$$

Thus, the eigenvalues of B are squared the eigenvalues of A, and the eigenvectors of B are identical to the eigenvectors of A. Then, we have $\lambda_B^2 = \lambda_A$, implying $\lambda_B = \sqrt{\lambda_A}$. This gives us our final result, where λ_i is the ith largest eigenvalue of A and v_i is the corresponding eigenvector.

$$\|v_i\|_2 = \frac{1}{\sqrt{\lambda_i}}$$

Thus, the eigenvalues of A determine the length along each eigenvector for an ellipsoid $x^T Ax$.

2 Singular Value Decomposition (SVD)