Alon Rosen

Eigenvalues and Variance Directions

Spectral methods extract directions along which a symmetric matrix acts most strongly. In this course we mainly use eigenvalues and eigenvectors to find directions of unusually large variance.

The main points of this chapter are:

- Symmetric matrices have orthonormal eigenvectors and real eigenvalues.
- The top eigenvector maximizes the quadratic form $\mathbf{v}^\top A\mathbf{v}$ over unit vectors.
- Rank-one matrices act in one direction and vanish on its orthogonal complement.

1 Eigenvalues of symmetric matrices

Let $A \in \mathbb{R}^{n\times n}$. A nonzero vector $\mathbf{v}$ is an *eigenvector* of A with *eigenvalue* λ if $A\mathbf{v} = \lambda\mathbf{v}$. We will only need the case in which A is symmetric.

Fact 1 (Spectral theorem)**.** *If $A = A^\top$, then there is an orthogonal matrix U and real numbers $\lambda_1, \dots, \lambda_n$ such that*

$$A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^\top.$$

Equivalently, A has an orthonormal basis of eigenvectors.

Writing the columns of U as $\mathbf{u}_1, \dots, \mathbf{u}_n$, every vector $\mathbf{v}$ can be written as $\mathbf{v} = \sum_i \alpha_i \mathbf{u}_i$, and then

$$A\mathbf{v} = \sum_i \lambda_i \alpha_i \mathbf{u}_i, \qquad \mathbf{v}^\top A\mathbf{v} = \sum_i \lambda_i \alpha_i^2.$$

Thus the eigendecomposition describes how A acts independently along orthogonal directions.

2 The top eigenvector

Assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. For a unit vector $\mathbf{v} = \sum_i \alpha_i \mathbf{u}_i$, we have $\sum_i \alpha_i^2 = 1$, so

$$\mathbf{v}^\top A\mathbf{v} = \sum_i \lambda_i \alpha_i^2 \leq \lambda_1.$$

Equality is attained at $\mathbf{v} = \mathbf{u}_1$.

Fact 2. *For a symmetric matrix A,*

$$\max_{\|\mathbf{v}\|_2 = 1} \mathbf{v}^\top A\mathbf{v} = \lambda_{\max}(A),$$

and any unit eigenvector corresponding to $\lambda_{\max}(A)$ attains the maximum.

This is the basic spectral optimization used later. If Σ is a covariance matrix, then $\mathbf{v}^\top \Sigma \mathbf{v}$ is the variance of $\mathbf{v}^\top \mathbf{X}$. Hence a top eigenvector of Σ is a direction of largest variance.

3 Rank-one matrices

Let $\mathbf{u} \in \mathbb{R}^n$. The matrix $\mathbf{u}\mathbf{u}^\top$ acts on a vector $\mathbf{v}$ by

$$(\mathbf{u}\mathbf{u}^\top)\mathbf{v} = \mathbf{u}(\mathbf{u}^\top \mathbf{v}).$$

Thus its output is always a multiple of $\mathbf{u}$.

Fact 3. *If $\mathbf{u} \neq 0$, then $\mathbf{u}\mathbf{u}^\top$ has rank one. Moreover, $\mathbf{u}$ is an eigenvector with eigenvalue $\|\mathbf{u}\|_2^2$, while every vector orthogonal to $\mathbf{u}$ is an eigenvector with eigenvalue 0.*

Indeed, $(\mathbf{u}\mathbf{u}^\top)\mathbf{u} = \|\mathbf{u}\|_2^2 \mathbf{u}$, while $(\mathbf{u}\mathbf{u}^\top)\mathbf{v} = 0$ whenever $\mathbf{u}^\top \mathbf{v} = 0$.

This is exactly what a rank-one covariance perturbation does: a term of the form $c\,\mathbf{u}\mathbf{u}^\top$ adds $c\|\mathbf{u}\|_2^2$ in the $\mathbf{u}$ direction and nothing in directions orthogonal to $\mathbf{u}$.