Matrix functions

1. We can easily define a natural *power* of a square matrix $\mathbf{X}$ as a product:

$$\mathbf{X}^n = \underbrace{\mathbf{XX...X}}_{n \text{ times}}$$

For a non-singular matrix, a negative power can be defined as a corresponding positive power of the matrix inverse:

$$\mathbf{X}^{-n} = (\mathbf{X}^{-1})^n$$

or, alternatively, as the matrix inverse of the positive power:

$$\mathbf{X}^{-n} = (\mathbf{X}^n)^{-1}$$

Exercise: Demonstrate that the two definitions are equivalent.

The zeroth power of any matrix is the identity matrix by definition:

$$\mathbf{X}^0 = \mathbf{I}$$

Exercise: Demonstrate the following properties of matrix powers:.

$$\mathbf{X}^n\mathbf{X}^m = \mathbf{X}^{n+m}$$

$$(\mathbf{X}^n)^m = \mathbf{X}^{nm}$$

3. If $\mathbf{y}$ is an eigenvector of matrix $\mathbf{X}$: $\mathbf{Xy} = \lambda\mathbf{y}$, then $\mathbf{X}^2\mathbf{y} = \mathbf{XXy} = \mathbf{X}\lambda\mathbf{y} = \lambda\mathbf{Xy} = \lambda^2\mathbf{y}$. Analogously, for any power: $\mathbf{X}^n\mathbf{y} = \lambda^n\mathbf{y}$. Therefore, the n-th power of matrix $\mathbf{X}$ has the same eigenvectors as the original matrix $\mathbf{X}$, with the corresponding eigenvalues being λ^n. This is also true for any negative power, including the matrix inverse: $\mathbf{X}^{-1}\mathbf{y} = \lambda\mathbf{y} \Rightarrow \mathbf{XX}^{-1}\mathbf{y} = \lambda\mathbf{Xy} \Rightarrow \mathbf{Xy} = (1/\lambda)\mathbf{y}$.

A *diagonalizable* matrix $\mathbf{X}$ can be expressed as follows (similarity transformation):

$$\mathbf{X} = \mathbf{T\Lambda T}^{-1}$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2,... \lambda_N)$ is the diagonal matrix of eigenvalues and $\mathbf{T}$ is the eigenvector matrix (column by column). Then, for the $\mathbf{X}^2$ matrix we obtain:

$$\mathbf{X}^2 = \mathbf{XX} = \left(\mathbf{T\Lambda T}^{-1}\right)\left(\mathbf{T\Lambda T}^{-1}\right) = \mathbf{T\Lambda}\left(\mathbf{T}^{-1}\mathbf{T}\right)\mathbf{\Lambda T}^{-1} = \mathbf{T\Lambda}^2\mathbf{T}^{-1}$$

where $\Lambda^2 = \text{diag}(\lambda_1^2, \lambda_2^2,... \lambda_N^2)$. In the case of normal matrices ($\mathbf{X}^\dagger\mathbf{X} = \mathbf{XX}^\dagger$), the eigenvectors *are* orthogonal: $\mathbf{T}^{-1} = \mathbf{T}^\dagger$. Analogously, any power $\mathbf{X}^n$ can be expressed as follows:

$$\mathbf{X}^n = \mathbf{T\Lambda}^n\mathbf{T}^{-1}$$

$$\Lambda^n = \text{diag}(\lambda_1^n, \lambda_2^n,... \lambda_N^n).$$

4. If a given real function can be expressed as a Taylor series:

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + ... = \sum_{k=0}^{\infty} a_k x^k$$

where

$$a_0 = f(0)$$

$$a_1 = \frac{f'(0)}{1!}$$

$$a_2 = \frac{f''(0)}{2!}$$

$$a_3 = \frac{f'''(0)}{3!}$$

$$\dots$$

we can construct a similar power series substituting the scalar argument x by a square matrix $\mathbf{X}$:

$$f(\mathbf{X}) = a_0 + a_1\mathbf{X} + a_2\mathbf{X}^2 + a_3\mathbf{X}^3 + ... = \sum_{k=0}^{\infty} a_k \mathbf{X}^k$$

This is a definition of function f of a matrix $\mathbf{X}$.

For instance, the *exponent* of matrix $\mathbf{X}$ will be given by the following series:

$$e^{\mathbf{X}} \equiv \exp(\mathbf{X}) = \mathbf{I} + \frac{1}{1!}\mathbf{X} + \frac{1}{2!}\mathbf{X}^2 + \frac{1}{3!}\mathbf{X}^3 + ... = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{X}^k$$

5. The above Taylor expansion can be used as such to *iteratively* compute a matrix function, but for *diagonalizable* matrices it is usually more efficient to employ the following algorithm:

- **a.** Diagonalize matrix $\mathbf{X}$, obtain the matrix of eigenvectors $\mathbf{T}$ and the diagonal matrix of eigenvalues Λ. If $\mathbf{X}$ is not a normal matrix, we will also need to invert $\mathbf{T}$. This gives a similarity transformation $\mathbf{X} = \mathbf{T}\Lambda\mathbf{T}^{-1}$, where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ... \lambda_N)$.
- **b.** The exponent of matrix $\mathbf{X}$ is given by the following expression:

$$\exp(\mathbf{X}) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{X}^k = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\mathbf{T}\boldsymbol{\Lambda}^k\mathbf{T}^{-1}\right) = \mathbf{T}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{\Lambda}^k\right)\mathbf{T}^{-1} =$$

$$= \mathbf{T}\mathrm{diag}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\lambda_1^k, \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_2^k, ... \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_N^k\right)\mathbf{T}^{-1} =$$

$$= \mathbf{T}\mathrm{diag}\left(\exp(\lambda_1), \exp(\lambda_2), ... \exp(\lambda_N)\right)\mathbf{T}^{-1}$$

  The latter formula gives us a non-iterative method of calculating $\exp(\mathbf{X})$ through diagonalization. In general, any matrix function can be calculated this way:
  $f(\mathbf{X}) = \mathbf{T}\ \mathrm{diag}(f(\lambda_1), f(\lambda_2), ... f(\lambda_N))\mathbf{T}^{-1}$.

Obviously, $f(\mathbf{X})$ has the same eigenvectors as the original matrix $\mathbf{X}$, with the corresponding eigenvalues being $f(\lambda)$.

6. The *square root* $\mathbf{X}^{1/2}$ of matrix $\mathbf{X}$ is defined through the expression $(\mathbf{X}^{1/2})^2 = \mathbf{X}$. The calculation of $\mathbf{X}^{1/2}$ can be also carried out by way of the diagonalization technique:

$$\mathbf{X}^{1/2} = \mathbf{T}\ \mathrm{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, ... \sqrt{\lambda_N}\right)\mathbf{T}^{-1}$$

Exercise: Demonstrate that the above formula indeed yields $\mathbf{X}^{1/2}$.

6. The *inverse square root* $\mathbf{X}^{-1/2}$ of matrix $\mathbf{X}$ is defined through the expression $(\mathbf{X}^{-1/2})^2 = \mathbf{X}^{-1}$. The calculation of $\mathbf{X}^{-1/2}$ can be done analogously:

$$\mathbf{X}^{-1/2} = \mathbf{T}\ \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, ... \frac{1}{\sqrt{\lambda_N}}\right)\mathbf{T}^{-1}$$

Of course, this is only possible if all eigenvalues are non-zero. Otherwise, $\mathbf{X}$ is singular and $\mathbf{X}^{-1/2}$ does not exist. Note that the functions $f(x) = \sqrt{x}$ or $f(x) = 1/\sqrt{x}$ can *not* be represented by Taylor series. Nevertheless, the above algorithms work.

The inverse square root plays a crucial role in the *Löwdin orthogonalization* method.

7. Note that if matrix $\mathbf{X}$ is Hermitian, all its powers $\mathbf{X}^n$ and matrix functions $f(\mathbf{X})$ will be also Hermitian. This refers, in particular, to matrix inverse $\mathbf{X}^{-1}$, matrix square root $\mathbf{X}^{1/2}$, and the inverse square root $\mathbf{X}^{-1/2}$.