Modified on Wednesday 2^{nd} December, 2020, 11:39:0411:39

Dec 2, 2020

- **Derivative of an eigenvalue**
- Consider a $n \times n$ matrix $A(t)$ that changes with the time t. Assume that $A = A(0)$ has distinct eigenvalues. Let $\lambda(t)$ be an eigenvalue of $A(t)$. In fact, there are n eigenvalues $\lambda_i(t)$ for $i = 1, 2, \ldots, n$. Let $\mathbf{x}(t)$ denote the eigenvector of $A(t)$ corresponding to eigenvalue $\lambda(t)$, $\mathbf{x}_i(t)$ for $\lambda_i(t)$.

$$X = X(t) = \begin{bmatrix} \mathbf{x}_1(t) & \cdots & \mathbf{x}_n(t) \end{bmatrix}, \quad D(t) = \text{diag}(\lambda_1(t), \ldots, \lambda_n(t))$$

$$A(t) = XDX^{-1}, \quad A^T(t) = (X^{-1})^T DX^T$$

Let $\mathbf{y}_i(t)$ denote the i-th column of $(X^{-1})^T$, which is an eigenvector of $A^T(t)$ corresponding to eigenvalue $\lambda_i(t)$.

$$A(t)\mathbf{x}(t) = \lambda(t)\mathbf{x}(t), \quad A^T(t)\mathbf{y}(t) = \lambda(t)\mathbf{y}(t), \quad \mathbf{y}^T(t)\mathbf{x}(t) = 1$$

So we have

$$\mathbf{y}^T(t)A(t)\mathbf{x}(t) = \lambda(t)$$

and

$$\begin{aligned}
\frac{d\lambda}{dt} &= \frac{d\mathbf{y}^T}{dt}A(t)\mathbf{x}(t) + \mathbf{y}^T(t)\frac{dA}{dt}\mathbf{x}(t) + \mathbf{y}^T(t)A(t)\frac{d\mathbf{x}}{dt} \\
&= \lambda(t)\frac{d\mathbf{y}^T}{dt}\mathbf{x}(t) + \mathbf{y}^T(t)\frac{dA}{dt}\mathbf{x}(t) + \lambda(t)\mathbf{y}^T(t)\frac{d\mathbf{x}}{dt} \\
&= \mathbf{y}^T(t)\frac{dA}{dt}\mathbf{x}(t) + \lambda(t)\left(\frac{d\mathbf{y}^T}{dt}\mathbf{x}(t) + \mathbf{y}^T(t)\frac{d\mathbf{x}}{dt}\right) \\
&= \mathbf{y}^T(t)\frac{dA}{dt}\mathbf{x}(t) + \lambda(t)\frac{d}{dt}(\mathbf{y}^T(t)\mathbf{x}(t)) \\
&= \mathbf{y}^T(t)\frac{dA}{dt}\mathbf{x}(t).
\end{aligned}$$

- Exercise. Let $A(t) = \begin{bmatrix} 1+2t & 1 \\ 2t & 2 \end{bmatrix}$, where $\lambda_1(0) = 2$ is an eigenvalue of $A(0)$. Find the derivative of eigenvalue $\lambda_1(t)$ at $t = 0$ for $A(t)$.
- **Derivative of a singular value**
- Assume that A has distinct singular values, $\sigma_i(t)$.

$$A(t)\mathbf{v}(t) = \sigma(t)\mathbf{u}(t), \quad A^T(t)\mathbf{u}(t) = \sigma(t)\mathbf{v}(t), \quad \mathbf{u}^T(t)\mathbf{u}(t) = \mathbf{v}^T(t)\mathbf{v}(t) = 1$$

$$A = U\Sigma V^T, \quad U^T AV = \Sigma, \quad \mathbf{u}^T(t)A(t)\mathbf{v}(t) = \mathbf{u}^T(t)\sigma(t)\mathbf{u}(t) = \sigma(t)$$

From $\sigma(t) = \mathbf{u}^T(t)A(t)\mathbf{v}(t)$, we obtain

$$\begin{aligned}
\frac{d\sigma}{dt} &= \frac{d\mathbf{u}^T}{dt}A(t)\mathbf{v}(t) + \mathbf{u}^T(t)\frac{dA}{dt}\mathbf{v}(t) + \mathbf{u}^T(t)A(t)\frac{d\mathbf{v}}{dt} \\
&= \sigma(t)\frac{d\mathbf{u}^T}{dt}\mathbf{u}(t) + \mathbf{u}^T(t)\frac{dA}{dt}\mathbf{v}(t) + \sigma(t)\mathbf{v}^T(t)\frac{d\mathbf{v}}{dt} \\
&= \sigma(t)\frac{d}{dt}\left(\frac{1}{2}\mathbf{u}^T(t)\mathbf{u}(t)\right) + \mathbf{u}^T(t)\frac{dA}{dt}\mathbf{v}(t) + \sigma(t)\frac{d}{dt}\left(\frac{1}{2}\mathbf{v}^T(t)\mathbf{v}(t)\right) \\
&= \mathbf{u}^T(t)\frac{dA}{dt}\mathbf{v}(t).
\end{aligned}$$

- Exercise. Let $A(t) = \begin{bmatrix} 1+2t & 1 \\ 2t & 2 \end{bmatrix}$. Find the derivative of singular value $\sigma_1(t)$ at $t = 0$ for $A(t)$.
- **Interlacing eigenvalues and rank one signals**
- How do eigenvalues and singular values change under perturbation with rank one matrices? Rank one perturbation can be a signal in application.
- Let S be a real symmetric matrix, of $n \times n$. We know that S has a spectral decomposition,

$$S = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T,$$

  with $\|\mathbf{v}_i\| = 1$ for all i and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

  Let $\mathbf{u}$ be a nonzero column vector in $\mathbb{R}^n$. Let z_i denote the eigenvalues of $S + \mathbf{u}\mathbf{u}^T$, $z_1 \geq z_2 \geq \cdots \geq z_n$. Then we have

$$z_1 \geq \lambda_1 \geq z_2 \geq \lambda_2 \geq \cdots \geq z_n \geq \lambda_n.$$

  The above inequalities mean that eigenvalues S and $S + \mathbf{u}\mathbf{u}^T$ are interlacing. This property can be easily proved for $\mathbf{u} = \alpha \mathbf{v}_i$.
- **A graphical explanation of interlacing of eigenvalues of $S + \theta \mathbf{u}\mathbf{u}^T$**
- For simplicity, we assume that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. Let z be an eigenvalue of $S + \theta \mathbf{u}\mathbf{u}^T$ with eigenvector $\mathbf{v}$. Note that z depends on θ and $\mathbf{u}$.

$$(S + \theta \mathbf{u}\mathbf{u}^T)\mathbf{v} = z\mathbf{v}$$

$$(zI - S)\mathbf{v} = \theta \mathbf{u}(\mathbf{u}^T \mathbf{v})$$

$$\mathbf{v} = (zI - S)^{-1} \theta \mathbf{u}(\mathbf{u}^T \mathbf{v})$$

  Multiply the both sides to the left with $\mathbf{u}^T$ and divide by a number $\theta \mathbf{u}^T \mathbf{v}$.

$$\mathbf{u}^T \mathbf{v} = \mathbf{u}^T (zI - S)^{-1} \theta \mathbf{u}(\mathbf{u}^T \mathbf{v})$$

$$\frac{1}{\theta} = \mathbf{u}^T (zI - S)^{-1} \mathbf{u}$$

  On the other hand, we have

$$(zI - S)\mathbf{v}_i = (z - \lambda_i)\mathbf{v}_i, \quad (zI - S)^{-1} \mathbf{v}_i = \frac{\mathbf{v}_i}{z - \lambda_i}.$$

  Express $\mathbf{u}$ as a linear combination of $\mathbf{v}_i$'s,

$$\mathbf{u} = \sum_{i=1}^{n} c_i \mathbf{v}_i.$$

  For simplicity, we assume that c_i's are nonzero.

$$\frac{1}{\theta} = \mathbf{u}^T (zI - S)^{-1} \mathbf{u} = \mathbf{u}^T (zI - S)^{-1} \sum_{i=1}^{n} c_i \mathbf{v}_i = \mathbf{u}^T \sum_{i=1}^{n} c_i (zI - S)^{-1} \mathbf{v}_i = \mathbf{u}^T \sum_{i=1}^{n} \frac{c_i \mathbf{v}_i}{z - \lambda_i} = \sum_{i=1}^{n} \frac{c_i^2}{z - \lambda_i}$$

$$\frac{1}{\theta} = \sum_{i=1}^{n} \frac{c_i^2}{z - \lambda_i}$$

  Draw on (z, y)-plane the two graphs:

$$y = y(z) = \sum_{i=1}^{n} \frac{c_i^2}{z - \lambda_i}, \quad y = \frac{1}{\theta}$$

  The intersections of two graphs give eigenvalues z_i's. The graph shows the interlacing property.

  We made some assumptions for simplicity, but we can explain interlacing property in full generality.