Quiz

- Give the SVD-based algorithm for solving least squares, and
- justify the algorithm by that showing it outputs the correct answer.
- Under what circumstances would this algorithm be preferred over the QR-based algorithm?

[12] The Eigenvector

Two interest-bearing accounts

Suppose Account 1 yields 5% interest and Account 2 yields 3% interest.

Represent balances in the two accounts by a 2-vector $\mathbf{x}^{(t)} = \begin{bmatrix} \text{amount in Account 1} \\ \text{amount in Account 2} \end{bmatrix}$.

$$\mathbf{x}^{(t+1)} = \begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix} \mathbf{x}^{(t)}$$

Let A denote the matrix. It is diagonal.

To find out how, say, $\mathbf{x}^{(100)}$ compares to $\mathbf{x}^{(0)}$, we can use Equation repeatedly:

$$\begin{aligned} \mathbf{x}^{(100)} &= A\mathbf{x}^{(99)} \\ &= A(A\mathbf{x}^{(98)}) \\ &\vdots \\ &= \underbrace{A \cdot A \cdot \cdots \cdot A}_{100 \text{ times}} \mathbf{x}^{(0)} \\ &= A^{100}\mathbf{x}^{(0)} \end{aligned}$$

Two interest-bearing accounts

$$
\begin{aligned}
\mathbf{x}^{(100)} &= A\mathbf{x}^{(99)} \\
&= A(A\mathbf{x}^{(98)}) \\
&\vdots \\
&= \underbrace{A \cdot A \cdot \cdots \cdot A}_{100 \text{ times}} \mathbf{x}^{(0)} \\
&= A^{100}\mathbf{x}^{(0)}
\end{aligned}
$$

Since A is a diagonal matrix, easy to compute powers of A:

Two interest-bearing accounts

$$
\begin{aligned}
\mathbf{x}^{(100)} &= A\mathbf{x}^{(99)} \\
&= A(A\mathbf{x}^{(98)}) \\
&\vdots \\
&= \underbrace{A \cdot A \cdot \cdots \cdot A}_{100 \text{ times}} \mathbf{x}^{(0)} \\
&= A^{100}\mathbf{x}^{(0)}
\end{aligned}
$$

Since A is a diagonal matrix, easy to compute powers of A:

$$
\begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix} \begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix} = \begin{bmatrix} 1.05^2 & 0 \\ 0 & 1.03^2 \end{bmatrix}
$$

Two interest-bearing accounts

$$\begin{aligned}
\mathbf{x}^{(100)} &= A\mathbf{x}^{(99)} \\
&= A(A\mathbf{x}^{(98)}) \\
&\vdots \\
&= \underbrace{A \cdot A \cdot \dots \cdot A}_{100 \text{ times}} \mathbf{x}^{(0)} \\
&= A^{100}\mathbf{x}^{(0)}
\end{aligned}$$

Since A is a diagonal matrix, easy to compute powers of A:

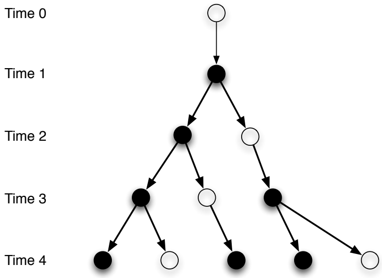
$$\underbrace{\begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix} \dots \begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix}}_{100 \text{ times}} = \begin{bmatrix} 1.05^{100} & 0 \\ 0 & 1.03^{100} \end{bmatrix} \approx \begin{bmatrix} 131.5 & 0 \\ 0 & 19.2 \end{bmatrix}$$

The takeaway: $\begin{bmatrix} \text{Account 1 balance after } t \text{ years} \\ \text{Account 2 balance after } t \text{ years} \end{bmatrix} = \begin{bmatrix} 1.05^t \cdot (\text{initial Account 1 balance}) \\ 1.03^t \cdot (\text{initial Account 2 balance}) \end{bmatrix}$

Rabbit reproduction

To avoid getting into trouble, I'll pretend sex doesn't exist.

- Each month, each adult rabbit gives birth to one baby.
- A rabbit takes one month to become an adult.
- Rabbits never die.



$$\begin{bmatrix} \text{adults at time } t+1 \\ \text{juveniles at time } t+1 \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \begin{bmatrix} \text{adults at time } t \\ \text{juveniles at time } t \end{bmatrix}$$

Use $\mathbf{x}^{(t)} = \begin{bmatrix} \text{number of adults after } t \text{ months} \\ \text{number of juveniles after } t \text{ months} \end{bmatrix}$

Then $\mathbf{x}^{(t+1)} = A\mathbf{x}^{(t)}$ where $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

$[1, 0],\ [1, 1],\ [2, 1],\ [3, 2],\ [5, 3],\ [8, 3], \ldots$

Analyzing rabbit reproduction

$\mathbf{x}^{(t+1)} = A\mathbf{x}^{(t)}$ where $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

As in bank-account example, $\mathbf{x}^{(t)} = A^t \mathbf{x}^{(0)}$.

Calculate how the entries of $\mathbf{x}^{(t)}$ grow as a function of t? With bank accounts, A was diagonal.

Not this time! However, there is a workaround:

Let $S = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$. Then $S^{-1}AS = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$.

$$\begin{aligned} A^t &= \underbrace{A\,A\,\cdots A}_{t \text{ times}} \\ &= (S\Lambda S^{-1})(S\Lambda S^{-1})\cdots(S\Lambda S^{-1}) \\ &= S\Lambda^t S^{-1} \end{aligned}$$

Λ is a diagonal matrix $\Rightarrow$ easy to compute Λ^t.

If $\Lambda = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}$ then $\Lambda^t = \begin{bmatrix} \lambda_1^t & \\ & \lambda_2^t \end{bmatrix}$. Here $\Lambda = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \\ & \frac{1-\sqrt{5}}{2} \end{bmatrix}$.

Interpretation using change of basis

Interpretation: To make the analysis easier, we will use a change of basis

Basis consists of the two columns of the matrix S, $\mathbf{v}_1 = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$

Let $\mathbf{u}^{(t)}$ = coordinate representation of $\mathbf{x}^{(t)}$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$.

- (`rep2vec`) To go from repres. $\mathbf{u}^{(t)}$ to vector $\mathbf{x}^{(t)}$ itself, we multiply $\mathbf{u}^{(t)}$ by S.
- (Move forward one month) To go from $\mathbf{x}^{(t)}$ to $\mathbf{x}^{(t+1)}$, we multiply $\mathbf{x}^{(t)}$ by A.
- (`vec2rep`) To go to coord. repres., we multiply by S^{-1}.

Multiplying by the matrix $S^{-1}AS$ carries out the three steps above.

But $S^{-1}AS = \Lambda = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$ so $\mathbf{u}^{(t+1)} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \mathbf{u}^{(t)}$

so

$$\mathbf{u}^{(t)} = \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^t & 0 \\ 0 & (\frac{1-\sqrt{5}}{2})^t \end{bmatrix} \mathbf{u}^{(0)}$$

Eigenvalues and eigenvectors

For this topic, consider only matrices A such that row-label set = col-label set (*endomorphic*).

> **Definition:** If λ is a scalar and $\mathbf{v}$ is a nonzero vector such that $A\mathbf{v} = \lambda\mathbf{v}$, we say that λ is an *eigenvalue* of A, and $\mathbf{v}$ is a corresponding *eigenvector*.

Convenient to require eigenvector has norm one.

Example: $\begin{bmatrix} 1.05 & 0 \\ 0 & 1.03 \end{bmatrix}$ has eigenvalues 1.05 and 1.03, and corresponding eigenvectors $[1, 0]$ and $[0, 1]$.

Example: $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ has eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$, and corresponding eigenvectors $[\frac{1+\sqrt{5}}{2}, 1]$ and $[\frac{1-\sqrt{5}}{2}, 1]$.

Example: What does it mean when A has 0 as an eigenvalue? There is a nonzero vector $\mathbf{v}$ such that $A\mathbf{v} = 0\mathbf{v}$. That is, A's null space is nontrivial.

Find an eigenvector corresp. to eigenvalue 0? Find nonzero vector in the null space.
What about other eigenvalues?

Eigenvector corresponding to an eigenvalue

Suppose λ is an eigenvalue of A, with corresponding eigenvector $\mathbf{v}$.
$A\mathbf{v} = \lambda\,\mathbf{v}$. $\Rightarrow A\mathbf{v} - \lambda\,\mathbf{v}$ is the zero vector.
$A\mathbf{v} - \lambda\,\mathbf{v} = (A - \lambda\,\mathbb{1})\mathbf{v}, \Rightarrow (A - \lambda\,\mathbb{1})\mathbf{v}$ is the zero vector.
That means that $\mathbf{v}$ is a nonzero vector in the null space of $A - \lambda\,\mathbb{1}$.
That means that $A - \lambda\,\mathbb{1}$ is not invertible.
Conversely, suppose $A - \lambda\,\mathbb{1}$ is not invertible
It is square, so it must have a nontrivial null space.
Let $\mathbf{v}$ be a nonzero vector in the null space.
Then $(A - \lambda\mathbb{1})\mathbf{v} = \mathbf{0}$, so $A\mathbf{v} = \lambda\mathbf{v}$.
We have proved the following:

Lemma: Let A be a square matrix.

- The number λ is an eigenvalue of A if and only if $A - \lambda\,\mathbb{1}$ is not invertible.
- If λ is in fact an eigenvalue of A then the corresponding eigenspace is the null space of $A - \lambda\,\mathbb{1}$.

Corollary

If λ is an eigenvalue of A then it is an eigenvalue of A^T.

Similarity

Definition: Two matrices A and B are *similar* if there is an invertible matrix S such that $S^{-1}AS = B$.

Proposition: Similar matrices have the same eigenvalues.

Proof: Suppose λ is an eigenvalue of A and $\mathbf{v}$ is a corresponding eigenvector. By definition, $A\mathbf{v} = \lambda\,\mathbf{v}$. Suppose $S^{-1}AS = B$, and let $\mathbf{w} = S^{-1}\mathbf{v}$. Then

$$
\begin{aligned}
B\mathbf{w} &= S^{-1}AS\mathbf{w} \\
&= S^{-1}ASS^{-1}\mathbf{v} \\
&= S^{-1}A\mathbf{v} \\
&= S^{-1}\lambda\,\mathbf{v} \\
&= \lambda\,S^{-1}\mathbf{v} \\
&= \lambda\,\mathbf{w}
\end{aligned}
$$

which shows that λ is an eigenvalue of B.

Example of similarity

Example: It is not hard to show that the eigenvalues of the matrix $A = \begin{bmatrix} 6 & 3 & -9 \\ 0 & 9 & 15 \\ 0 & 0 & 15 \end{bmatrix}$ are its diagonal elements (6, 9, and 15) because A is upper triangular. The matrix $B = \begin{bmatrix} 92 & -32 & -15 \\ -64 & 34 & 39 \\ 176 & -68 & -99 \end{bmatrix}$ has the property that $B = S^{-1}AS$ where $S = \begin{bmatrix} -2 & 1 & 4 \\ 1 & -2 & 1 \\ -4 & 3 & 5 \end{bmatrix}$.
Therefore the eigenvalues of B are also 6, 9, and 15.

Diagonalizability

Definition: If A is similar to a diagonal matrix,we say A is *diagonalizable*.
(if there is an invertible matrix S such that $S^{-1}AS = \Lambda$ where Λ is a diagonal matrix)

Equation $S^{-1}AS = \Lambda$ is equivalent to equation $A = S\Lambda S^{-1}$, which is the form used in the analysis of rabbit population. How is diagonalizability related to eigenvalues?

- Eigenvalues of a diagonal matrix $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ are its diagonal entries.
- If matrix A is similar to Λ then the eigenvalues of A are the eigenvalues of Λ
- Equation $S^{-1}AS = \Lambda$ is equivalent to $AS = S\Lambda$. Write S in terms of columns:

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \left[\begin{array}{c|c|c} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{array} \right] = \left[\begin{array}{c|c|c} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{array} \right] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Diagonalizability

Definition: If A is similar to a diagonal matrix,we say A is *diagonalizable*.
(if there is an invertible matrix S such that $S^{-1}AS = \Lambda$ where Λ is a diagonal matrix)

Equation $S^{-1}AS = \Lambda$ is equivalent to equation $A = S\Lambda S^{-1}$, which is the form used in the analysis of rabbit population. How is diagonalizability related to eigenvalues?

- Eigenvalues of a diagonal matrix $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ are its diagonal entries.
- If matrix A is similar to Λ then the eigenvalues of A are the eigenvalues of Λ
- Equation $S^{-1}AS = \Lambda$ is equivalent to $AS = S\Lambda$. Write S in terms of columns:

$$\left[\begin{array}{c|c|c} A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{array}\right] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Diagonalizability

Definition: If A is similar to a diagonal matrix,we say A is *diagonalizable*.
(if there is an invertible matrix S such that $S^{-1}AS = \Lambda$ where Λ is a diagonal matrix)

Equation $S^{-1}AS = \Lambda$ is equivalent to equation $A = S\Lambda S^{-1}$, which is the form used in the analysis of rabbit population. How is diagonalizability related to eigenvalues?

- Eigenvalues of a diagonal matrix $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ are its diagonal entries.
- If matrix A is similar to Λ then the eigenvalues of A are the eigenvalues of Λ
- Equation $S^{-1}AS = \Lambda$ is equivalent to $AS = S\Lambda$. Write S in terms of columns:

$$\left[\begin{array}{c|c|c} A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \end{array} \right] = \left[\begin{array}{c|c|c} \lambda_1\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n \end{array} \right]$$

  Columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ of S are eigenvectors. S is invertible $\Rightarrow$ eigenvectors lin. indep.
- The argument goes both ways: if $n \times n$ matrix A has n linearly independent eigenvectors then A is diagonalizable.

Diagonalizability Theorem

Diagonalizability Theorem: An $n \times n$ matrix A is diagonalizable iff it has n linearly independent eigenvectors.

Example: Consider the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Its null space is trivial so zero is not an eigenvalue.

For any 2-vector $\begin{bmatrix} x \\ y \end{bmatrix}$, we have $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ y \end{bmatrix}$.

Suppose λ is an eigenvector. Then for some vector $[x, y]$,

$$\lambda \;\; [x, y] = [x + y, y]$$

Therefore $\lambda\, y = y$. Therefore $y = 0$. Therefore every eigenvector is in Span $\{[1, 0]\}$. Thus the matrix does not have two linearly independent eigenvectors, so it is not diagonalizable.

Interpretation using change of basis, re-revisited

Suppose $n \times n$ matrix A is diagonalizable, so it has linearly independent e-vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ with e-values are $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Any vector $\mathbf{x}$ can be written as a linear combination:

$$\mathbf{x} = \alpha_1 \, \mathbf{v}_1 + \cdots + \alpha_n \, \mathbf{v}_n$$

Left-multiply by A on both sides of the equation:

$$\begin{aligned} A\mathbf{x} &= A(\alpha_1 \mathbf{v}_1) + A(\alpha_2 \mathbf{v}_2) + \cdots + A(\alpha_n \mathbf{v}_n) \\ &= \alpha_1 A\mathbf{v}_1 + \alpha_2 A\mathbf{v}_2 + \cdots + \alpha_n A\mathbf{v}_n \\ &= \alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_2 \mathbf{v}_2 + \cdots + \alpha_n \lambda_n \mathbf{v}_n \end{aligned}$$

Applying the same reasoning to $A(A\mathbf{x})$, we get

$$A^2\mathbf{x} = \alpha_1 \lambda_1^2 \mathbf{v}_1 + \alpha_2 \lambda_2^2 \mathbf{v}_2 + \cdots + \alpha_n \lambda_n^2 \mathbf{v}_n$$

More generally, for any nonnegative integer t,

$$A^t\mathbf{x} = \alpha_1 \lambda_1^t \mathbf{v}_1 + \alpha_2 \lambda_2^t \mathbf{v}_2 + \cdots + \alpha_n \lambda_n^t \mathbf{v}_n$$

If $|\lambda_1| > |\lambda_2|$ then eventually λ_1^t will be *much* bigger than $\lambda_2^t, \ldots, \lambda_n^t$, so first term will dominate. For a large enough value of t, $A^t\mathbf{x}$ will be approximately $\alpha_1 \lambda_1^t \mathbf{v}_1$.