Linear Control and Estimation

Eigenvalues and Eigenvectors

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

References

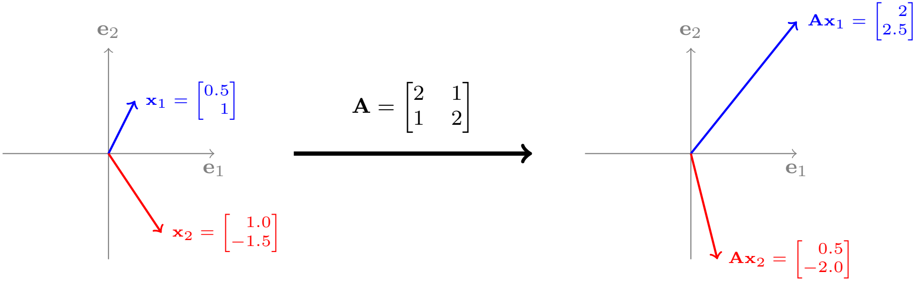
- G Strang, Linear Algebra: Chapters 5.

Linear transformation

- Matrices represent linear transformations, $\mathbf{A} \in \mathbb{R}^{m \times n}$ represents a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$.

$$\mathbf{y} = T(\mathbf{x}) = \mathbf{A}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{y} \in \mathbb{R}^m$$

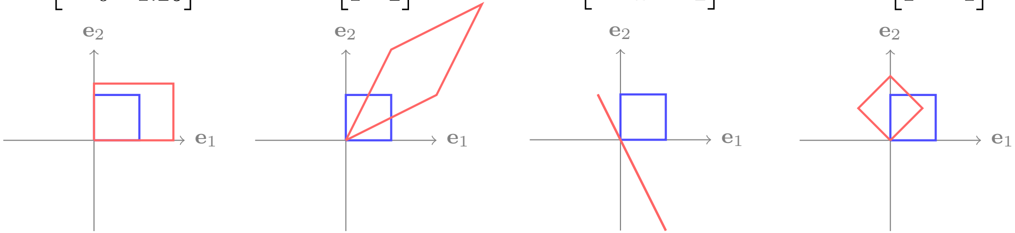
- Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$.
  In general, T scales and rotates the vector $\mathbf{x}$ to produce $\mathbf{y}$.



Linear transformation

An easier way is to look at what happens to the standard basis $\{\mathbf{e}_i\}_{i=1}^{n}$.

$$\mathbf{A} = \begin{bmatrix} 1.75 & 0 \\ 0 & 1.25 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} -0.5 & 1 \\ 1 & -2 \end{bmatrix} \qquad \mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$



Complex Vectors and Matrics

- Similar to $\mathbb{R}^n$, we can have $\mathbb{C}^n$. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{r1} + jx_{i1} \\ x_{r2} + jx_{i2} \\ \vdots \\ x_{rn} + jx_{in} \end{bmatrix}$
- Vector addition and scalar mulitplication are the same. The scalar is a complex number.
- Additive identity, and scalar multiplication identity are the same. So is the **standard basis** $\{\mathrm{e}_i\}_{i=1}^n$
- **Linear independence**: The set $\{\mathbf{v}_i\}_{i=1}^n$ with $\mathbf{v}_i \in \mathbb{C}^n$ is linearly independent, if $\sum_{i=1}^n c_i \mathbf{v}_i = 0, \implies c_i = 0, \; \forall 1 \leq i \leq n, \;\; c_i \in \mathbb{C}^n$
- **Inner product**: $\mathbf{x}^H \mathbf{y} = \begin{bmatrix} \overline{x}_1 & \overline{x}_2 & \dots & \overline{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n \overline{x}_i y_i$

Complex Vectors and Matrics

- **Length**: $\|\mathbf{x}\|_2^2 = \mathbf{x}^H\mathbf{x} = \sum_{i=1}^n \overline{x}_i x_i = \sum_{i=1}^n |x_i|^2$
- **Orthogonality**: $\mathbf{x}^H\mathbf{y} = 0$
- Complex matrices have complex entries. $\mathbf{A} \in \mathbb{C}^{m\times n}$ such that $a_{ij} \in \mathbb{C},\ \forall 1 \leq i \leq m,\ 1 \leq j \leq n$
- The transpose operation is generalized to conjugate transpose known as the *Hermitian*. $\mathbf{A}^H = \overline{\mathbf{A}}^T$.
- The idea of symmertric matrices $\mathbb{R}^{n\times n}$ are now generalized to $\mathbb{C}^{n\times n}$ as $\mathbf{A} = \mathbf{A}^H$. Such matrices are called **Hermitian** matrices.
- Orthogonal matrices in the complex case are called **Unitary** matrices, $\mathbf{U}^H\mathbf{U} = \mathbf{I} \implies \mathbf{U}^{-1} = \mathbf{U}^H$.

Eignenvectors and Eigenvalues

- It turns out for any linear transformation represented by $\mathbf{A} \in \mathbb{R}^{n \times n}$, the are vectors that satisfy with the following property,

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}, \quad \mathbf{x} \in \mathbb{C}^n, \lambda \in \mathbb{C}$$

Least Squares Methods

- The general solution to this least squares problem can be derived using calculus. Let $f(x) = \|b - Ax\|$

$$\nabla f(x) = 0 \longrightarrow \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 0$$

  Going through the algebra, we end up with the following expression for $\hat{x}$ that minimizes $f(x)$,

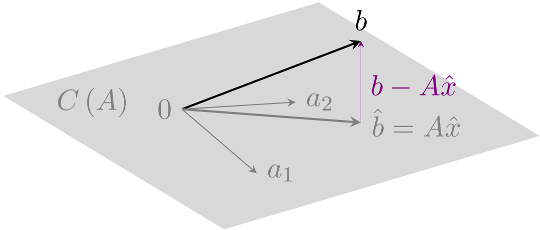
$$\underbrace{A^T A \hat{x} = A^T b}_{\text{Normal Equations}}$$

  A is full rank, $\implies A^T A$ is invertible.

$$\implies \hat{x} = \underbrace{\left(A^T A\right)^{-1} A^T}_{\text{Pseudo-inverse}} b = A^\dagger b$$

Least Squares Methods

- $\hat{x}$ is the approximate least squares solution. $\hat{b} = A\hat{x}$, which is in general not equal to b. *When is* $b = \hat{b}$*?*
- We know two things about $\hat{b}$,
  1. $\hat{b} \in C(A)$: $\hat{b}$ is the column space of A.
  2. $\left\| b - \hat{b} \right\|$ is minimum.



$$\|b - A\hat{x}\|_2^2 \text{ is minimum } \implies (b - A\hat{x}) \perp A\hat{x}$$

$$(A\hat{x})^T (b - A\hat{x}) = 0 \implies \hat{x}^T \underbrace{\left(A^T b - A^T A\hat{x}\right)}_{\text{Normal Equations}} = 0$$

The least squares approximate solution of $Ax = b$ is the solution solution to the equation $Ax = \hat{b}$, where $\hat{b}$ is the projection of b onto the column space of A $(C(A))$

Multi-Objective Least Squares

- There are applications where there is more than one objective that must be optimized,
$$J_1 = \|A_1 x - b_1\|^2, \; J_2 = \|A_2 x - b_2\|^2, \; \ldots \; J_k = \|A_k x - b_k\|^2,$$
and often these are conflicting objectives.
- We can define a single objective function J that is takes into account the different objective functions.
$$J = \sum_{i=1}^{k} \rho_i J_i, \;\; \rho_i > 0, \;\; \forall 1 \leq i \leq k$$
- The ρ_is indicate the relative weightage given to the individual objectives.
$$J = J_1 + \sum_{i=2}^{k} \rho_i J_i$$

Multi-Objective Least Squares

$$J = \rho_1 \left\|A_1 x - b_1\right\|^2 + \ldots + \rho_k \left\|A_k x - b_k\right\|^2 = \left\|\sqrt{\rho_1} A_1 x - \sqrt{\rho_1} b_1\right\|^2 + \ldots + \rho_k \left\|\sqrt{\rho_k} A_k x - \sqrt{\rho_k} b_k\right\|^2$$

$$J = \left\| \begin{bmatrix} \sqrt{\rho_1} A_1 \\ \sqrt{\rho_2} A_1 \\ \vdots \\ \sqrt{\rho_k} A_k \end{bmatrix} x - \begin{bmatrix} \sqrt{\rho_1} b_1 \\ \sqrt{\rho_2} b_1 \\ \vdots \\ \sqrt{\rho_k} b_k \end{bmatrix} \right\|^2 = \left\| \tilde{A} x - \tilde{b} \right\|^2 \implies \hat{x} = \left(\tilde{A}^T \tilde{A} \right)^{-1} \tilde{A}^T \tilde{b}$$

The columns of $\tilde{A}$ are must be independent, which happens if the columns of at least one of the A_is is independent.

Consider a two objective case, $J = J_1 + \rho J_2$.

$$\hat{x} = \begin{cases} \operatorname{argmin}_x \left\|A_1 x - b_1\right\|^2 & \rho = 0 \\ \operatorname{argmin}_x \left\|A_2 x - b_2\right\|^2 & \rho \to \infty \end{cases}$$

Multi-Objective Least Squares

Any solution that lies on this curve is called the *Pareto optimal* solution. There exists no solution $\tilde{x}$, such that $J_1(\tilde{x}) \leq J_1(\hat{x})$ and $J_2(\tilde{x}) \leq J_2(\hat{x})$ where, both inqualities hold strictly.

Multi-Objective Least Squares

- Multi-objective least squares methods play an important role in both control and estimation problems.
- Appropriate choice of the objective functions can also help deal with conditions where the columns of A_i are not independent. Consider the following example,

$$J_1 = \|A_1 x - b_1\|^2 \quad \text{and} \quad J_2 = \|A_2 x - b - 2\|^2$$

  where, $A_1 \in \mathbb{R}^{m_1 \times n}$ and $A_2 \in \mathbb{R}^{m_2 \times n}$, such that $m_1, m_2 < n$. Thus, the columns of A_1 and A_2 are not independent! However, if $m_1 + m_2 \geq n$, then it is possible that the columns of $\tilde{A}$ are independent.
- This is called **regularized least squares**.
- **Tichonov regularization**: $J = \|Ax - y\|^2 + \rho \|x\|^2$, where $\rho > 0$.

$$\tilde{A} = \begin{bmatrix} A \\ \sqrt{\rho} I \end{bmatrix} \implies \hat{x} = \left(A^T A + \rho I\right)^{-1} A^T b$$

Multi-Objective Least Squares

- **Tichonov regularization**: $J = \|Ax - y\|^2 + \rho\|x\|^2$, where $\rho > 0$.

$$\tilde{A} = \begin{bmatrix} A \\ \sqrt{\rho}I \end{bmatrix} \implies \hat{x} = \left(A^T A + \rho I\right)^{-1} A^T b$$

- $\hat{x}$ gives a unique solution in minimizing J, even when A is not full rank.
- Even when A is full rank, the regularization term can be used to improve the condition number of the matrix.

Multi-Objective Least Squares

- **Tichonov regularization**: $J = \|Ax - y\|^2 + \rho \|x\|^2$, where $\rho > 0$.

$$\tilde{A} = \begin{bmatrix} A \\ \sqrt{\rho} I \end{bmatrix} \implies \hat{x} = \left(A^T A + \rho I\right)^{-1} A^T b$$

- $\hat{x}$ gives a unique solution in minimizing J, even when A is not full rank.
- Even when A is full rank, the regularization term can be used to improve the condition number of the matrix.

Constrained Least Squares

- **Problem**:
$$\begin{aligned} \text{minimize} \quad & \|Ax - b\|^2 \\ \text{subject to} \quad & Cx = d \end{aligned}$$
where, $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{p \times n}$ and $d \in \mathbb{R}^p$.
- This can be solved using the *method of Lagrange multipliers*. When we do this, we finally arrive the following set of equations, called the *Karush-Kuhn-Tucker* (KKT) equation,
$$2\left(A^T A\right)\hat{x} - 2A^T b + C^T \hat{z} = 0$$
$$\begin{bmatrix} 2\left(A^T A\right) & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 2A^T b \\ d \end{bmatrix}$$
- The coefficient matrix on the LHS of the KKT equation a square matrix of dimensions $(n+p) \times (n+p)$ is invertible, if and only if, $\begin{bmatrix} A \\ C \end{bmatrix}$ is full rank.