STATS 100C: Linear Models	**Spring 2017**

Lecture 1: April 4

Lecturer: Arash Amini *Scribes: Eric Chuu*

1.1 Review of Linear Algebra

Some topics that we should be familiar with:

- Fundamental subspaces with matrix $\mathrm{X} \in \mathbb{R}^{n\times p}$
- Rank-Nullity Theorem: The image of X: $[\mathrm{Im}(\mathrm{X})]^{\perp} = \ker(\mathrm{X}^T) \subseteq \mathbb{R}^n$
- Orthogonal Decomposition of a space with respect to a subspace $V \subseteq \mathbb{R}^n$ is given by

$$\mathbb{R}^n = V \oplus V^{\perp} \qquad \text{(concept of projection)}$$

- Spectral Decomposition: A symmetric matrix $\mathrm{A} \in \mathbb{R}^{n\times m}$ is given by

$$A = V\Lambda V^T$$

  where V is orthogonal and Λ is diagonal (eigenvalue decomposition)

Definition 1.1. *A set of vectors* $\{\mathbf{x_1}, \dots, \mathbf{x_n}\}$ *is linearly independent if*

$$\sum_{i=1}^{n} c_i \mathbf{x_i} = \mathbf{0} \quad \Rightarrow \quad c_i = 0, \text{ for } i = 1, \dots n$$

Exercise. Show that if $\{\mathbf{x_1}, \dots, \mathbf{x_n}\}$ is pairwise orthogonal, then they are linearly independent.

Definition 1.2. *The span of a set of vectors* $\{\mathbf{x_1}, \dots, \mathbf{x_n}\}$*, each in* $\mathbb{R}^d$ *is the set of all linear combinations of them.*

$$\text{span}\left(\{\mathbf{x_1}, \dots, \mathbf{x_n}\}\right) = \left\{\sum_{i=1}^{n} c_i \mathbf{x_i} : c_i \in \mathbb{R}\right\}$$

Note that the span of a set of vectors is a linear subspace of $\mathbb{R}^d$

Example.

$$\text{span}\left(\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\2\\0\end{bmatrix}\right\}\right) = \left\{t\begin{bmatrix}1\\1\\0\end{bmatrix} : t \in \mathbb{R}\right\} \tag{1.1}$$

$$c_1\begin{bmatrix}1\\1\\0\end{bmatrix} + c_2\begin{bmatrix}1\\1\\0\end{bmatrix} = (c_1 + 2c_2)\begin{bmatrix}1\\1\\0\end{bmatrix}, \quad t := c_1 + 2c_2 \tag{1.2}$$

Example.

$$\text{span}\left(\left\{\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}\right)=\left\{t_1\begin{bmatrix}1\\1\\0\end{bmatrix}+t_2\begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}=\left\{\begin{bmatrix}t_1+t_2\\t_1-t_2\\0\end{bmatrix}:t_1,t_2\in\mathbb{R}\right\} \tag{1.3}$$

Definition 1.3. *A basis of a subspace V is a set of linearly independent vectors whose span is V.*

Example. Using the previous example, we can construct a basis for $\mathbb{R}^2\times\{\mathbf{0}\}$

$$\text{span}\left(\left\{\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}\right)=\text{span}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}\right\}\right)=\left\{\begin{bmatrix}\alpha\\\beta\\0\end{bmatrix}:\alpha,\beta\in\mathbb{R}\right\} \tag{1.4}$$

Theorem 1.4. *All the bases for a subspace have the same number of elements. This number is called the dimension of the subspace (Dimension Theorem).*

Definition 1.5. *Consider the matrix $\mathrm{X}\in\mathbb{R}^{n\times p}$, which we can denote*

$$\mathrm{X}=\begin{bmatrix}|&|&&|\\\mathbf{x}_1&\mathbf{x}_2&\dots&\mathbf{x}_p\\|&|&&|\end{bmatrix}$$

Then the column space/range/image is given by

$$\begin{aligned}\mathrm{Im}(\mathrm{X}) &:= \text{span}\left(\{\mathbf{x_1},\dots,\mathbf{x_p}\}\right)\subseteq\mathbb{R}^n\\ &=\left\{\sum_{i=1}^{p}\beta_i\mathbf{x}_i:\beta_i\in\mathbb{R}\right\}\end{aligned}$$

Note that we can write the sum as

$$\sum_{i=1}^{p}\beta_i\mathbf{x}_i=\begin{bmatrix}|&|&&|\\\mathbf{x}_1&\mathbf{x}_2&\dots&\mathbf{x}_p\\|&|&&|\end{bmatrix}\begin{bmatrix}\beta_1\\\vdots\\\beta_p\end{bmatrix}=\mathrm{X}\boldsymbol{\beta}$$

Then the image of X can be written as $\mathrm{Im}(X)=\{\mathrm{X}\boldsymbol{\beta}:\boldsymbol{\beta}\in\mathbb{R}^p\}$, where X can be interpreted as an operator that maps $\mathbb{R}^p$ to some subset of $\mathbb{R}^n$. The image of a matrix X is a linear subspace of $\mathbb{R}^n$.

Definition 1.6. *The rank of a matrix is given by:*

$$\text{rank}(\mathrm{X})=\dim(\mathrm{Im}(\mathrm{X}))=\dim\left(\text{span}\left(\{\mathbf{x_1},\dots,\mathbf{x_p}\}\right)\right) \tag{1.5}$$

Definition 1.7. *The kernel or nullspace of a matrix* X *is the set*

$$\ker(\mathrm{X})=\{\boldsymbol{\beta}:\mathrm{X}\boldsymbol{\beta}=\mathbf{0}\} \tag{1.6}$$

$\ker(\mathrm{X})$ *is a linear subspace of* $\mathbb{R}^p$, *and the nullity of* $\mathrm{X}=\dim(\ker(\mathrm{X}))$

Example. Given the matrix

$$\mathrm{X} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

$$\mathrm{Im}(\mathrm{X}) = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} : t \in \mathbb{R} \right\} \quad \mathrm{rank}(\mathrm{X}) = 1$$

$$\begin{aligned} \ker(\mathrm{X}) &= \left\{ \boldsymbol{\beta} = (\beta_1, \beta_2) : \beta_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = \mathbf{0} \right\} \\ &= \left\{ \boldsymbol{\beta} = (\beta_1, \beta_2) : \beta_1 + 2\beta_2 = 0 \right\} \\ &= \mathrm{span}\left(\left\{ \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} \right\} \right), \quad \text{so } \dim(\ker(\mathrm{X})) = 1 \end{aligned}$$

□

1.1.1 Inner Products, Orthogonality

Definition 1.8. *For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the inner product of $\mathbf{x}, \mathbf{y}$ is defined as*

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} x_i y_i = \mathbf{x}^T \mathbf{y} \tag{1.7}$$

Definition 1.9. *We say that $\mathbf{x}$ is orthogonal to $\mathbf{y}$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. We say that the vector $\mathbf{x}$ is orthogonal to a subspace $V \subseteq \mathbb{R}^n$, $\mathbf{x} \perp V$, if and only if $\mathbf{x} \perp \mathbf{y}, \forall y \in V$.*

Definition 1.10. *The set of all vectors $\mathbf{x} \perp V$ is called the orthogonal complement of V, denoted $V^\perp$.*

$$V^\perp = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} \perp V \right\} \tag{1.8}$$

The orthogonal complement forms a linear subspace.