Math S-21b – Lecture #15 Notes

In this class we'll apply the **Spectral Theorem** and the **Principal Axes Theorem** to understand **quadratic forms**. We'll also discuss the **Singular Value Decomposition** of any matrix.

Definition: A quadratic form is a homogeneous polynomial of degree 2, i.e. a polynomial function $q(\mathbf{x})$ such that $q(t\mathbf{x}) = t^2\mathbf{x}$, a pure quadratic expression in n variables.. For example:

(a) $q(x,y) = 8x^2 - 4xy + 5y^2$ (b) $q(x_1,x_2,x_3) = x_1{}^2 - 2x_2{}^2 + 4x_3{}^2 + 2x_1x_2 - x_1x_3 + 4x_2x_3$

Observation: Any quadratic form can be expressed as $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ where $\mathbf{x}$ is an $n\times 1$ (column) vector and **A** is a symmetric $n\times n$ matrix. For example:

(a) $q(x,y) = 8x^2 - 4xy + 5y^2 = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$

(b) $q(x_1,x_2,x_3) = x_1{}^2 - 2x_2{}^2 + 4x_3{}^2 + 2x_1x_2 - x_1x_3 + 4x_2x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & 1 & -\frac{1}{2} \\ 1 & -2 & 2 \\ -\frac{1}{2} & 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$

Principal Axes Theorem: Any quadratic form $q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ may be expressed without cross terms in new coordinates via an orthonormal change of basis. That is, there exists an orthonormal basis $\mathcal{B} = \{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ and scalars $\{\lambda_1,\cdots,\lambda_n\}$ such that if $\{y_1,\cdots,y_n\}$ are the coordinates relative to the basis $\mathcal{B} = \{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$, then $q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y} = \lambda_1 y_1{}^2 + \cdots + \lambda_n y_n{}^2$.

Proof: By the Spectral Theorem, since **A** is symmetric it is orthogonally diagonalizable, i.e. it has real eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$ and associated orthonormal eigenvectors $\mathcal{B} = \{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ for the matrix **A** such that if is the orthogonal change of basis matrix $\mathbf{S} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{u}_1 & \cdots & \mathbf{u}_n \\ \downarrow & & \downarrow \end{bmatrix}$, then $[\mathbf{A}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{S}^{\mathrm{T}}\mathbf{A}\mathbf{S} = \mathbf{D} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, a diagonal matrix. We may therefore write $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{\mathrm{T}}$, so
$q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{S}\mathbf{D}\mathbf{S}^{\mathrm{T}}\mathbf{x} = (\mathbf{S}^{\mathrm{T}}\mathbf{x})^{\mathrm{T}}\mathbf{D}(\mathbf{S}^{\mathrm{T}}\mathbf{x}) = (\mathbf{S}^{-1}\mathbf{x})^{\mathrm{T}}\mathbf{D}(\mathbf{S}^{-1}\mathbf{x}) = \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y} = \lambda_1 y_1{}^2 + \cdots + \lambda_n y_n{}^2$.

Definitions: A quadratic form $q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ with symmetric matrix **A** is called:

(a) **positive definite** if all of the eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$ are strictly positive.
(b) **negative definite** if all of the eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$ are strictly negative.
(c) **positive semi-definite** if all of the eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$ are greater than or equal to 0.
(d) **negative semi-definite** if all of the eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$ are less than or equal to 0.
(e) **indefinite** if some eigenvalues are positive and some or negative (and some may be equal to 0).

Some immediate applications are in **identifying the graphs of quadratic functions** $q(x,y)$ and **identifying level sets** of the form $q(x,y) = C$ (conic sections such as ellipses and hyperbolas) and $q(x,y,z) = C$ (quadric sections such as ellipsoids and hyperboloids of one or two sheets).

Example: The graph of the function $q(x,y)=8x^2-4xy+5y^2=\begin{bmatrix} x & y\end{bmatrix}\begin{bmatrix} 8 & -2 \\ -2 & 5\end{bmatrix}\begin{bmatrix} x \\ y\end{bmatrix}=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ can be easily identified by calculating the eigenvalues of **A** as $\lambda_1=9$ and $\lambda_2=4$ with orthonormal eigenbasis $\mathbf{u}_1=\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1\end{bmatrix}$, $\mathbf{u}_2=\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2\end{bmatrix}$. In (rotated) coordinates $\{u,v\}$ relative to this new basis, the graph $q=9u^2+4v^2$ can be identified as an (upward, rotated) **paraboloid**. Similarly, the level set $q(x,y)=8x^2-4xy+5y^2=36$ may be re-expressed as $q=9u^2+4v^2=36$ or $\left(\frac{u}{2}\right)^2+\left(\frac{v}{3}\right)^2=1$, a (rotated) **ellipse** with semi-major axis 3 and semi-minor axis 2. Note that the longer axis corresponds to the *smaller* eigenvalue (slower growth) and the shorter axis corresponds to the *larger* eigenvalue (faster growth) to reach the given level set.

In the case of a level set of a **quadratic function in 3 variables**, if $q(x_1,x_2,x_3)=C>0$. If the signs of the eigenvalues are $\{+,+,+\}$, the level set will be a (rotated) **ellipsoid**. If the signs of the eigenvalues are $\{+,+,-\}$, the level set will be a (rotated) **hyperboloid of one sheet**. If the signs of the eigenvalues are $\{+,-,-\}$, the level set will be a (rotated) **hyperboloid of two sheets**.

Singular Values and the Singular Value Decomposition (SVD)

Given any $m\times n$ matrix **A**, it's possible to find an orthonormal basis $\mathcal{B}=\{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ for the domain ($\mathbf{R}^n$) as well as an orthonormal basis $\mathcal{C}=\{\mathbf{w}_1,\cdots,\mathbf{w}_m\}$ for the codomain ($\mathbf{R}^m$) such that the images $\{\mathbf{A}\mathbf{u}_1,\cdots,\mathbf{A}\mathbf{u}_n\}$ are orthogonal (some may be $\mathbf{0}$) and are scalar multiples, respectively, of the vectors $\{\mathbf{w}_1,\cdots,\mathbf{w}_m\}$ in the codomain. The scalars $\|\mathbf{A}\mathbf{u}_i\|=\sigma_i$ are called the **singular values** of the matrix **A** (and the linear transformation that it represents).

This observation follows by considering the symmetric $n\times n$ matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. By the Spectral Theorem, this matrix yields an orthonormal basis of eigenvectors $\mathcal{B}=\{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ with real eigenvalues $\{\lambda_1,\cdots,\lambda_n\}$.
Furthermore, $(\mathbf{A}\mathbf{u}_i)\cdot(\mathbf{A}\mathbf{u}_j)=\langle\mathbf{A}\mathbf{u}_i,\mathbf{A}\mathbf{u}_j\rangle=(\mathbf{A}\mathbf{u}_i)^{\mathrm{T}}\mathbf{A}\mathbf{u}_j=\mathbf{u}_i^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u}_j=\lambda_j\mathbf{u}_i^{\mathrm{T}}\mathbf{u}_j=\lambda_j\langle\mathbf{u}_i,\mathbf{u}_j\rangle$.
In the case where $i\neq j$, this yields that $\langle\mathbf{A}\mathbf{u}_i,\mathbf{A}\mathbf{u}_j\rangle=0$, i.e. that these images are orthogonal (or one or both could be $\mathbf{0}$). In the case where $i=j$, this yields that $\langle\mathbf{A}\mathbf{u}_j,\mathbf{A}\mathbf{u}_j\rangle=\|\mathbf{A}\mathbf{u}_j\|^2=\lambda_j$, so all of these eigenvalues must be greater than or equal to 0. Furthermore, $\|\mathbf{A}\mathbf{u}_j\|=\sqrt{\lambda_j}=\sigma_j$ are the singular values. If we order the eigenvalues (and the singular values) in decreasing order, and if we create the orthogonal $n\times n$ matrix $\mathbf{P}=\begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{u}_1 & \cdots & \mathbf{u}_n \\ \downarrow & & \downarrow\end{bmatrix}$ from the orthonormal basis for the domain and the $m\times m$ matrix $\mathbf{Q}=\begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{w}_1 & \cdots & \mathbf{w}_m \\ \downarrow & & \downarrow\end{bmatrix}$ from the orthonormal basis for the codomain, we can create the following commuting diagram describing of vectors are transformed relative to the standard bases and these preferred orthonormal bases:

$$\begin{array}{ccc} \{\mathbf{R}^n,\mathcal{E}_n\} & \xrightarrow{\mathbf{A}} & \{\mathbf{R}^m,\mathcal{E}_m\} \\ \mathbf{P}\uparrow & & \mathbf{Q}\uparrow \\ \{\mathbf{R}^n,\mathcal{B}\} & \xrightarrow{\Sigma} & \{\mathbf{R}^m,\mathcal{C}\} \end{array}$$

Since **P** and **Q** are orthogonal matrices, $\mathbf{P}^{-1}=\mathbf{P}^{\mathrm{T}}$ and $\mathbf{Q}^{-1}=\mathbf{Q}^{\mathrm{T}}$, so we get $\mathbf{A}=\mathbf{Q}\boldsymbol{\Sigma}\mathbf{P}^{\mathrm{T}}$. This is known as the **Singular Value Decomposition** (SVD). If $m=n$, i,e, if A is a square matrix, then the matrix $\boldsymbol{\Sigma}$ will be diagonal with the singular values on the diagonal. However the SVD also applies to any matrix – in which case the singular values will still appear along the "diagonal" starting at the upper-left position with 0's everywhere else.

Example: Consider the matrix $\mathbf{A}=\begin{bmatrix}6 & 2\\ -7 & 6\end{bmatrix}$. We calculate $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\begin{bmatrix}6 & -7\\ 2 & 6\end{bmatrix}\begin{bmatrix}6 & 2\\ -7 & 6\end{bmatrix}=\begin{bmatrix}85 & -30\\ -30 & 40\end{bmatrix}=\mathbf{B}$. This yields the eigenvalues (in decreasing order) $\lambda_1=100$ and $\lambda_2=25$. These yield, respectively, the orthonormal basis vectors $\mathbf{u}_1=\frac{1}{\sqrt{5}}\begin{bmatrix}2\\ -1\end{bmatrix}$, $\mathbf{u}_2=\frac{1}{\sqrt{5}}\begin{bmatrix}1\\ 2\end{bmatrix}$. We then calculate $\mathbf{Au}_1=10\left(\frac{1}{\sqrt{5}}\begin{bmatrix}1\\ -2\end{bmatrix}\right)=10\mathbf{w}_1$, $\mathbf{Au}_2=5\left(\frac{1}{\sqrt{5}}\begin{bmatrix}2\\ 1\end{bmatrix}\right)=5\mathbf{w}_2$. This yields the orthogonal matrices $\mathbf{P}=\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ -1 & 2\end{bmatrix}$ and $\mathbf{Q}=\frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\ -2 & 1\end{bmatrix}$ and $\mathbf{\Sigma}=\begin{bmatrix}10 & 0\\ 0 & 5\end{bmatrix}$.

Example: Consider the matrix $\mathbf{A}=\begin{bmatrix}0 & 1 & 1\\ 1 & 1 & 0\end{bmatrix}$. We calculate $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\begin{bmatrix}0 & 1\\ 1 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & 1 & 1\\ 1 & 1 & 0\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 1\end{bmatrix}=\mathbf{B}$. This yields the eigenvalues (in decreasing order) $\lambda_1=3$, $\lambda_2=1$, and $\lambda_3=0$. These yield, respectively, the orthonormal basis vectors $\mathbf{u}_1=\frac{1}{\sqrt{6}}\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}$, $\mathbf{u}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 0\\ -1\end{bmatrix}$, and $\mathbf{u}_3=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\ -1\\ 1\end{bmatrix}$. We then calculate $\mathbf{Au}_1=\sqrt{3}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}\right)=\sqrt{3}\mathbf{w}_1$,

$\mathbf{Au}_2=1\left(\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\ 1\end{bmatrix}\right)=1\mathbf{w}_2$, and $\mathbf{Au}_3=\mathbf{0}$. This yields the orthogonal matrices $\mathbf{P}=\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3}\\ 2/\sqrt{6} & 0 & -1/\sqrt{3}\\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3}\end{bmatrix}$ and

$\mathbf{Q}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$ and $\mathbf{\Sigma}=\begin{bmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\end{bmatrix}=\begin{bmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\end{bmatrix}$.

So $\mathbf{A}=\mathbf{Q\Sigma P}^{\mathrm{T}}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6}\\ 1/\sqrt{2} & 0 & -1/\sqrt{2}\\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3}\end{bmatrix}$ is the corresponding singular value decomposition.

Notes by Robert Winters