Math S-21b – Lecture #9 Notes

The main topics in this lecture are orthogonal projection, the Gram-Schmidt orthogonalization process, QR factorization, isometries and orthogonal transformations, least-squares approximate solutions and applications to data-fitting.

Some previous results:

1) Suppose $V = \text{Span}\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$. Let $\mathbf{A} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_k \\ \downarrow & & \downarrow \end{bmatrix}$. This is an $n \times k$ matrix with $V = \text{im}(\mathbf{A})$ and $\boxed{V^\perp = (\text{im}\,\mathbf{A})^\perp = \ker(\mathbf{A}^\text{T})}$.

2) Suppose $\mathscr{B} = \{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ is an orthonormal (ON) basis for a subspace $V \subseteq \mathbf{R}^n$. Then for any $\mathbf{x} \in \mathbf{R}^n$, $\boxed{\text{Proj}_V \mathbf{x} = (\mathbf{x}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{x}\cdot\mathbf{u}_k)\mathbf{u}_k}$. If we write $\mathbf{B} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{u}_1 & \cdots & \mathbf{u}_k \\ \downarrow & & \downarrow \end{bmatrix}$, then $\boxed{\text{Proj}_V = \mathbf{B}\mathbf{B}^\text{T}}$ is the matrix for orthogonal projection onto *V*, and $\boxed{\text{Ref}_V = 2\mathbf{B}\mathbf{B}^\text{T} - \mathbf{I}}$ is the matrix for reflection through this subspace.

3) If $V = \mathbf{R}^n$ and $\mathscr{B} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ is an orthonormal basis for all of $\mathbf{R}^n$, then $\mathbf{B} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{u}_1 & \cdots & \mathbf{u}_n \\ \downarrow & & \downarrow \end{bmatrix}$ will be an $n \times n$ matrix with ON columns (hence invertible), and $\text{Proj}_V = \mathbf{B}\mathbf{B}^\text{T} = \mathbf{I}$. Therefore <u>in this special case</u> we'll have $\mathbf{B}^{-1} = \mathbf{B}^\text{T}$. Such a matrix is called an **orthogonal matrix**.

4) If $\mathbf{B} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{u}_1 & \cdots & \mathbf{u}_k \\ \downarrow & & \downarrow \end{bmatrix}$ is any $n \times k$ matrix with orthonormal columns, then $\mathbf{B}^\text{T}\mathbf{B} = \mathbf{I}_k$. In the special case where $\mathbf{B}$ is an $n \times n$ matrix with orthonormal columns, this gives $\mathbf{B}^\text{T}\mathbf{B} = \mathbf{I}_n$.

Transpose Facts

The following relations hold wherever the expressions are defined:

(1) $(\mathbf{AB})^\text{T} = \mathbf{B}^\text{T}\mathbf{A}^\text{T}$

(2) $(\mathbf{A}+\mathbf{B})^\text{T} = \mathbf{A}^\text{T} + \mathbf{B}^\text{T}$

(3) If $\mathbf{A}$ is an invertible $n \times n$ matrix, then $\mathbf{A}^\text{T}$ is also invertible and $(\mathbf{A}^\text{T})^{-1} = (\mathbf{A}^{-1})^\text{T}$

The proofs are somewhat routine. For example, to establish (1), if $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is a $n \times p$ matrix, then the (i, j) of $\mathbf{AB}$ will be $\sum_{k=1}^{n} a_{ik}b_{kj}$. This will then be the (j,i) entry of $(\mathbf{AB})^\text{T}$. On the other hand, the (j,k) entry of $\mathbf{B}^\text{T}$ will be b_{kj} and the (k,i) entry of $\mathbf{A}^\text{T}$ will be a_{ik}, so the (j,i) entry of $\mathbf{B}^\text{T}\mathbf{A}^\text{T}$ will be $\sum_{k=1}^{n} b_{kj}a_{ik} = \sum_{k=1}^{n} a_{ik}b_{kj}$ which coincides with the (j,i) entry of $(\mathbf{AB})^\text{T}$. Therefore $(\mathbf{AB})^\text{T} = \mathbf{B}^\text{T}\mathbf{A}^\text{T}$.

Corollary: The matrix $\mathbf{A}$ for any orthogonal projection or reflection is always symmetric, i.e. $\mathbf{A}^\text{T} = \mathbf{A}$.

Proof: Using the previous results, any projection matrix can be expressed as $\mathbf{A} = \mathbf{B}\mathbf{B}^\text{T}$ and $\mathbf{A}^\text{T} = (\mathbf{B}\mathbf{B}^\text{T})^\text{T} = \mathbf{B}\mathbf{B}^\text{T} = \mathbf{A}$, so the matrix is symmetric. Similarly, $\text{Ref}_V = 2\mathbf{B}\mathbf{B}^\text{T} - \mathbf{I}$ and $(2\mathbf{B}\mathbf{B}^\text{T} - \mathbf{I})^\text{T} = 2(\mathbf{B}\mathbf{B}^\text{T})^\text{T} - \mathbf{I}^\text{T} = 2\mathbf{B}\mathbf{B}^\text{T} - \mathbf{I}$, so this matrix is also symmetric.

Gram-Schmidt Orthogonalization Process

Suppose we begin with a basis $\mathscr{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ for a k-dimensional subspace $V \subseteq \mathbf{R}^n$. We would like to construct an orthonormal basis for this same subspace. The Gram-Schmidt orthogonalization process sequentially constructs such a basis. It should be emphasized that the resulting ON basis is very much dependent on the ordering of the original basis. We proceed as follows:

(1) Start with $\mathbf{v}_1$ and normalize it by scaling, i.e. $\mathbf{u}_1 = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$. For reasons that will soon become clear, we write $r_{11} = \|\mathbf{v}_1\|$. We can also solve for $\mathbf{v}_1 = r_{11}\mathbf{u}_1$. Let $V_1 = \text{Span}\{\mathbf{v}_1\} = \text{Span}\{\mathbf{u}_1\}$.

(2) Next, we take the second basis vector $\mathbf{v}_2$, find its projection onto the subspace V_1, subtract this from the original to get a vector orthogonal to the first, then scale this to get a unit vector. We can calculate the projection as $\text{Proj}_{V_1}(\mathbf{v}_2) = (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$, so we take $\mathbf{u}_2 = \dfrac{\mathbf{v}_2 - \text{Proj}_{V_1}(\mathbf{v}_2)}{\|\mathbf{v}_2 - \text{Proj}_{V_1}(\mathbf{v}_2)\|}$. Note that $r_{22} = \|\mathbf{v}_2 - \text{Proj}_{V_1}(\mathbf{v}_2)\|$ is the perpendicular height of the parallelogram determined by the vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ and the area of this parallelogram is therefore $(base)(\perp\, height) = r_{11}r_{22}$. We can also solve for $\mathbf{v}_2 = (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 + r_{22}\mathbf{u}_2$. Let $V_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

(3) If $k > 2$, we continue with the third basis vector $\mathbf{v}_3$. We find its projection onto the subspace V_2, subtract this from the original to get a vector orthogonal to V_2, then scale this to get a unit vector. We can calculate the projection as $\text{Proj}_{V_2}(\mathbf{v}_3) = (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2$, so we take $\mathbf{u}_3 = \dfrac{\mathbf{v}_3 - \text{Proj}_{V_2}(\mathbf{v}_3)}{\|\mathbf{v}_3 - \text{Proj}_{V_2}(\mathbf{v}_3)\|}$. Note that $r_{33} = \|\mathbf{v}_3 - \text{Proj}_{V_2}(\mathbf{v}_3)\|$ is the perpendicular height of the parallelepiped determined by the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and the volume of this parallelepiped is therefore $(area\ of\ base)(\perp\, height) = r_{11}r_{22}r_{33}$. We can also solve for $\mathbf{v}_3 = (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2 + r_{33}\mathbf{u}_3$. Let $V_3 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

We continue in this same manner until we exhaust our finite list of basis vectors. The last orthonormal vector will be $\mathbf{u}_k = \dfrac{\mathbf{v}_k - \text{Proj}_{V_{k-1}}(\mathbf{v}_k)}{\|\mathbf{v}_k - \text{Proj}_{V_{k-1}}(\mathbf{v}_k)\|}$ and if we write $r_{kk} = \|\mathbf{v}_k - \text{Proj}_{V_{k-1}}(\mathbf{v}_k)\|$ we can define the k-volume of the k-dimensional parallelepiped determined by the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ as $r_{11}r_{22}\cdots r_{kk}$. We can also solve for $\mathbf{v}_k = (\mathbf{v}_k \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{v}_k \cdot \mathbf{u}_{k-1})\mathbf{u}_{k-1} + r_{kk}\mathbf{u}_k$. We then have $V = \text{Span}\{\mathbf{v}_1, \cdots, \mathbf{v}_k\} = \text{Span}\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$, and this completes the orthogonalization process.

QR factorization

If we assemble the equations from the above process as $\left\{\begin{array}{l} \mathbf{v}_1 = r_{11}\mathbf{u}_1 \\ \mathbf{v}_2 = (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 + r_{22}\mathbf{u}_2 \\ \mathbf{v}_3 = (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2 + r_{33}\mathbf{u}_3 \\ \quad\vdots \\ \mathbf{v}_k = (\mathbf{v}_k \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{v}_k \cdot \mathbf{u}_{k-1})\mathbf{u}_{k-1} + r_{kk}\mathbf{u}_k \end{array}\right\}$ we can express this as a product of matrices as follows:

$$\mathbf{A}=\underbrace{\begin{bmatrix}\uparrow & \uparrow & & \uparrow\\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k\\ \downarrow & \downarrow & & \downarrow\end{bmatrix}}_{\substack{n\times k\text{ matrix w/linearly}\\ \text{independent columns}}}=\underbrace{\begin{bmatrix}\uparrow & \uparrow & & \uparrow\\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k\\ \downarrow & \downarrow & & \downarrow\end{bmatrix}}_{\substack{n\times k\text{ matrix}\\ \text{w/orthonormal columns}}}\underbrace{\begin{bmatrix}r_{11} & \mathbf{v}_2\cdot\mathbf{u}_1 & \cdots & \mathbf{v}_k\cdot\mathbf{u}_1\\ 0 & r_{22} & \cdots & \mathbf{v}_k\cdot\mathbf{u}_2\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & r_{kk}\end{bmatrix}}_{\substack{k\times k\text{ upper triangular matrix}\\ \text{with nonzero diagonal entries}}}=\mathbf{QR}$$

The columns of the matrix **A** are the original basis vectors; the columns of the matrix **Q** are those of the Gram-Schmidt basis; and the entries of the matrix **R** capture all of the geometric aspects of the original basis, i.e. lengths, areas, etc. and the non-orthogonality of the original vectors. Note that the <u>k-volume</u> is just the product of the diagonal entries of **R**, i.e. $r_{11}r_{22}\cdots r_{kk}$.

Example: In $\mathbf{R}^4$, let $\mathbf{v}_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$, $\mathbf{v}_2=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$, and $\mathbf{v}_3=\begin{bmatrix}0\\2\\1\\-1\end{bmatrix}$, and let $V=\text{Span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$. These vector form a basis for V, but not an orthonormal basis. Using the Gram-Schmidt process, we have $r_{11}=\|\mathbf{v}_1\|=2$, so $\mathbf{u}_1=\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$. We next calculate

$$\mathbf{v}_2-\text{Proj}_{V_1}(\mathbf{v}_2)=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}-\left(\frac{1}{2}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\right)\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}-\frac{1}{4}(2)\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}-\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}=\begin{bmatrix}1/2\\-1/2\\-1/2\\1/2\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}.$$

Its magnitude is $r_{22}=\|\mathbf{v}_2-\text{Proj}_{V_1}(\mathbf{v}_2)\|=1$, so $\mathbf{u}_2=\frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}$. We next calculate

$$\mathbf{v}_3-\text{Proj}_{V_2}(\mathbf{v}_3)=\mathbf{v}_3-(\mathbf{v}_3\cdot\mathbf{u}_1)\mathbf{u}_1-(\mathbf{v}_3\cdot\mathbf{u}_2)\mathbf{u}_2=\begin{bmatrix}0\\2\\1\\-1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}=\begin{bmatrix}1/2\\1/2\\-1/2\\-1/2\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}\text{, and}$$

$r_{33}=\|\mathbf{v}_3-\text{Proj}_{V_2}(\mathbf{v}_3)\|=1$, so $\mathbf{u}_3=\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$.

The 3-volume of the parallelepiped determined by $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ is $r_{11}r_{22}r_{33}=(2)(1)(1)=2$.

The corresponding QR-factorization is $\mathbf{A}=\begin{bmatrix}1&1&0\\1&0&2\\1&0&1\\1&1&-1\end{bmatrix}=\begin{bmatrix}1/2&1/2&1/2\\1/2&-1/2&1/2\\1/2&-1/2&-1/2\\1/2&1/2&-1/2\end{bmatrix}\begin{bmatrix}2&1&1\\0&1&-2\\0&0&1\end{bmatrix}=\mathbf{QR}$.

Isometries and orthogonal transformations

Given two spaces V and W where there's a notion of distance (metric spaces), an isometry is a transformation $T:V\to W$ that preserves distances. Familiar examples include rotations and reflections, but also "isometric embeddings" such as the transformation that places $\mathbf{R}^2$ in $\mathbf{R}^3$ as either the *xy*-plane, *xz*-plane, *yz*-plane, or any other plane such that distances are preserved. In the case of linear transformations, we are more specific:

Definition: A linear transformation $T:\mathbf{R}^n\to\mathbf{R}^n$ is called an **orthogonal transformation** if it preserves norms, i.e. $\|T(\mathbf{x})\|=\|\mathbf{x}\|$ for all **x**. Its matrix is called an **orthogonal matrix**.

Proposition: If a linear transformation $T:\mathbf{R}^n \to \mathbf{R}^m$ preserves norm, then $\ker(T)=\{\mathbf{0}\}$.
Proof: If $T(\mathbf{x})=\mathbf{0}$, then $\|T(\mathbf{x})\|=\|\mathbf{x}\|=\|\mathbf{0}\|=0$, so $\mathbf{x}=\mathbf{0}$.

Corollary: If $T:\mathbf{R}^n \to \mathbf{R}^n$ is an orthogonal transformation, it must be invertible.

Proposition: If $T:\mathbf{R}^n \to \mathbf{R}^n$ is an orthogonal transformation, then T preserves dot products: $T(\mathbf{x})\cdot T(\mathbf{y})=\mathbf{x}\cdot\mathbf{y}$ for all $\mathbf{x},\mathbf{y}\in\mathbf{R}^n$.
Proof: By linearity, $T(\mathbf{x}+\mathbf{y})=T(\mathbf{x})+T(\mathbf{y})$, so $\|T(\mathbf{x}+\mathbf{y})\|=\|T(\mathbf{x})+T(\mathbf{y})\|$ and $\|T(\mathbf{x}+\mathbf{y})\|^2=\|T(\mathbf{x})+T(\mathbf{y})\|^2$.
Since T is an orthogonal transformation,
$\|T(\mathbf{x}+\mathbf{y})\|^2=\|\mathbf{x}+\mathbf{y}\|^2=(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y})=\mathbf{x}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2\mathbf{x}\cdot\mathbf{y}$. Similarly,
$\|T(\mathbf{x})+T(\mathbf{y})\|^2=\|T(\mathbf{x})\|^2+\|T(\mathbf{y})\|^2+2T(\mathbf{x})\cdot T(\mathbf{y})=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2T(\mathbf{x})\cdot T(\mathbf{y})$. Comparing both sides we see that $T(\mathbf{x})\cdot T(\mathbf{y})=\mathbf{x}\cdot\mathbf{y}$.

Proposition: If $T:\mathbf{R}^n \to \mathbf{R}^n$ is an orthogonal transformation, then T preserves angles. That is, if θ_1 is the angle between two nonzero vectors **x** and **y**, and if θ_2 is the angle between $T(\mathbf{x})$ and $T(\mathbf{y})$, then $\theta_2=\pm\theta_1$.
Proof: We know that $\mathbf{x}\cdot\mathbf{y}=\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta_1$ and $T(\mathbf{x})\cdot T(\mathbf{y})=\|T(\mathbf{x})\|\|T(\mathbf{y})\|\cos\theta_2=\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta_2$, and $T(\mathbf{x})\cdot T(\mathbf{y})=\mathbf{x}\cdot\mathbf{y}$. Therefore $\cos\theta_1=\cos\theta_2$, so $\theta_2=\pm\theta_1$.

Matrix of an orthogonal transformation
Because the standard basis $\mathcal{E}=\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ is an orthonormal basis of $\mathbf{R}^n$ and since orthogonal transformations preserve length and angle, it follows that $\{T(\mathbf{e}_1),T(\mathbf{e}_2),\ldots,T(\mathbf{e}_n)\}$ must also be an orthonormal basis of $\mathbf{R}^n$. This includes rotations and reflections. The matrix of an orthogonal transformation must therefore be $\mathbf{A}=\begin{bmatrix}\uparrow & & \uparrow\\ [T(\mathbf{e}_1)]_{\mathcal{E}} & \cdots & [T(\mathbf{e}_n)]_{\mathcal{E}}\\ \downarrow & & \downarrow\end{bmatrix}=\begin{bmatrix}\uparrow & & \uparrow\\ \mathbf{A}\mathbf{e}_1 & \cdots & \mathbf{A}\mathbf{e}_n\\ \downarrow & & \downarrow\end{bmatrix}=\begin{bmatrix}\uparrow & & \uparrow\\ \mathbf{u}_1 & \cdots & \mathbf{u}_n\\ \downarrow & & \downarrow\end{bmatrix}$, i.e. it must have <u>orthonormal columns</u>. It must also be the case that $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\begin{bmatrix}\leftarrow & \mathbf{u}_1 & \rightarrow\\ & \vdots & \\ \leftarrow & \mathbf{u}_n & \rightarrow\end{bmatrix}\begin{bmatrix}\uparrow & & \uparrow\\ \mathbf{u}_1 & \cdots & \mathbf{u}_n\\ \downarrow & & \downarrow\end{bmatrix}=\begin{bmatrix}\mathbf{u}_1\cdot\mathbf{u}_1 & \cdots & \mathbf{u}_1\cdot\mathbf{u}_n\\ \vdots & \ddots & \vdots\\ \mathbf{u}_n\cdot\mathbf{u}_1 & \cdots & \mathbf{u}_n\cdot\mathbf{u}_n\end{bmatrix}=\begin{bmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1\end{bmatrix}=\mathbf{I}_n$, so an orthogonal matrix has the special property that $\mathbf{A}^{\mathrm{T}}=\mathbf{A}^{-1}$, and any matrix that satisfies this property must be the matrix of an orthogonal transformation. Geometrically, these are all (compositions of) rotations and reflections.

Least-Squares approximate solutions
Situation: We would like to solve a linear system $\mathbf{A}\mathbf{x}=\mathbf{b}$ where **A** is an $m\times n$ matrix, but we find that the system is <u>inconsistent</u>. This means that $\mathbf{b}\notin \operatorname{im}\mathbf{A}$, but this suggests the possibility that we might seek a vector $\mathbf{x}^*$ such that $\mathbf{A}\mathbf{x}^*$ is as close to the subspace $\operatorname{im}\mathbf{A}$ as possible. Orthogonal projection is a natural choice, so we seek $\mathbf{x}^*$ such that $\boxed{\mathbf{A}\mathbf{x}^*=\operatorname{Proj}_V\mathbf{b}}$ where $V=\operatorname{im}\mathbf{A}$. This means that we want $\mathbf{b}-\mathbf{A}\mathbf{x}^*\in(\operatorname{im}\mathbf{A})^{\perp}=V^{\perp}$. We have already shown that $(\operatorname{im}\mathbf{A})^{\perp}=\ker(\mathbf{A}^{\mathrm{T}})$, so we want $\mathbf{b}-\mathbf{A}\mathbf{x}^*\in\ker(\mathbf{A}^{\mathrm{T}})$, i.e. $\mathbf{A}^{\mathrm{T}}\left(\mathbf{b}-\mathbf{A}\mathbf{x}^*\right)=\mathbf{0}$ or $\boxed{\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}^*=\mathbf{A}^{\mathrm{T}}\mathbf{b}}$. This is known as the **normal equation** (or normal equations). A solution $\mathbf{x}^*$ is called a **least-squares approximate solution**.

The name "least-squares solution" comes from an alternate way that it can be derived using multivariable calculus methods in the special case where we're trying to find the line that best fits a given data set. That method involves minimizing the sum of the square deviations between values predicted by a best-fit line (also called a regression line) and actual values provided by the data set.

The normal equation is easy to remember. If the original system is $\mathbf{Ax}=\mathbf{b}$, then you just have to apply the matrix $\mathbf{A}^{\mathrm{T}}$ to both sides of the equation to get $\mathbf{A}^{\mathrm{T}}\mathbf{Ax}=\mathbf{A}^{\mathrm{T}}\mathbf{b}$. This system will always be consistent. If **A** is an $m\times n$ matrix, then $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ will be an $n\times n$ (square) matrix. It will also be symmetric since $(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}\mathbf{A}$.

In the case where $\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})=\{\mathbf{0}\}$, the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ will be invertible and there will be a unique least-squares solution $\mathbf{x}^*=(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$. Many students memorize this formula and apply it blindly, but it is often simplest to solve the consistent system $\mathbf{A}^{\mathrm{T}}\mathbf{Ax}=\mathbf{A}^{\mathrm{T}}\mathbf{b}$ using row reduction to find the least-squares solution.

There is a simple way to determine when the normal equation will yield a unique least-squares solution. This is based on the following lemma:

Lemma: For any matrix **A**, it is the case that $\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})=\ker\mathbf{A}$.

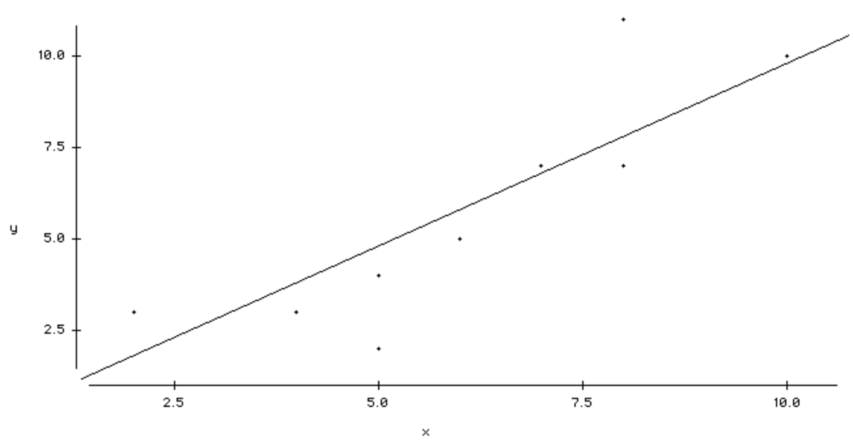
Proof: If $\mathbf{x}\in\ker\mathbf{A}$, then $\mathbf{Ax}=\mathbf{0}$. So $\mathbf{A}^{\mathrm{T}}\mathbf{Ax}=\mathbf{A}^{\mathrm{T}}\mathbf{0}=\mathbf{0}$ which means that $\mathbf{x}\in\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})$. So $\ker\mathbf{A}\subseteq\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})$. On the other hand, if $\mathbf{x}\in\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})$, then $\mathbf{A}^{\mathrm{T}}\mathbf{Ax}=\mathbf{0}$. But this means that $\mathbf{Ax}\in\ker(\mathbf{A}^{\mathrm{T}})=(\mathrm{im}\,\mathbf{A})^{\perp}$. But it's obvious that $\mathbf{Ax}\in\mathrm{im}\,\mathbf{A}$, so we have $\mathbf{Ax}\in(\mathrm{im}\,\mathbf{A})^{\perp}\cap(\mathrm{im}\,\mathbf{A})=\{\mathbf{0}\}$. Therefore $\mathbf{Ax}=\mathbf{0}$, and therefore $\mathbf{x}\in\ker\mathbf{A}$. So $\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})\subseteq\ker\mathbf{A}$. Therefore $\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A})=\ker\mathbf{A}$.

We also know that for any matrix **A**, $\ker\mathbf{A}=\{\mathbf{0}\}$ if and only if the columns of **A** are linearly independent. If we combine this fact and the previous results, we see that the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ will be invertible and there will be a unique least-squares approximate solution to $\mathbf{Ax}=\mathbf{b}$ if and only if the columns of **A** are linearly independent.

There's an unexpected benefit provided by the least-squares solution. If V is any subspace with basis $\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$, if we let $\mathbf{A}=\begin{bmatrix}\uparrow & & \uparrow\\ \mathbf{v}_1 & \cdots & \mathbf{v}_k\\ \downarrow & & \downarrow\end{bmatrix}$, then $V=\mathrm{im}\,\mathbf{A}$ and **A** will have linearly independent columns. So for any $\mathbf{b}\in\mathbf{R}^n$, $\mathrm{Proj}_V\mathbf{b}=\mathbf{Ax}^*=\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$. Therefore $\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}$ will be the matrix for orthogonal projection onto the subspace *V*. This is significant in that our previous method required the use of the Gram-Schmidt process to produce an orthonormal basis for the subspace *V*. This alternative method only requires that $\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$ be a basis. It is perhaps worth noting that if $\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$ had been an orthonormal basis, then we would have $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}_k$ and $\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}=\mathbf{AIA}^{\mathrm{T}}=\mathbf{AA}^{\mathrm{T}}$ which coincides with our previous method.

Data fitting

It is common that data occurs in the form of ordered pairs (or ordered *n*-tuples). If we plot the data, the resulting graph is called a scatterplot. If the scatterplot suggests a roughly straight-line relationship, it is reasonable to ask which straight line might best fit the given data.



Suppose the data is $\{(x_i,y_i)\}_{i=1}^{N}$. We can use our least-squares method by *assuming the absurd*, namely that all of the data fits a straight with equation $y=mx+b$ perfectly. If this is the case, then we get the system of linear equations:

$$\begin{Bmatrix} mx_1+b=y_1\\ mx_2+b=y_2\\ \vdots\\ mx_N+b=y_N\end{Bmatrix}\Rightarrow\begin{bmatrix}x_1 & 1\\ x_2 & 1\\ \vdots & \vdots\\ x_N & 1\end{bmatrix}\begin{bmatrix}m\\ b\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_N\end{bmatrix}\Rightarrow\mathbf{Ac}=\mathbf{y}$$

This is, of course, a hopelessly inconsistent linear system, but we can find a least-squares approximate solution by solving $\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{c} = \mathbf{A}^{\mathrm{T}}\mathbf{y}$. We can calculate $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & N \end{bmatrix}$ and

$\mathbf{A}^{\mathrm{T}}\mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} y_i \end{bmatrix}$, so the normal equations are $\begin{bmatrix} \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & N \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} y_i \end{bmatrix}$. These can then be easily solved to find the slope m and the intercept b for the lie of best fit.

Best quadratic?

It may be the case that the scatterplot suggests something other than a straight line relationship. If, for example, you suspect a quadratic relationship, start by writing this as $y = ax^2 + bx + c$. If we again assume the absurd possibility that all the data fits this quadratic perfectly, we get the system of linear equations:

$$\left\{ \begin{array}{c} ax_1^2 + bx_1 + c = y_1 \\ ax_2^2 + bx_2 + c = y_2 \\ \vdots \\ ax_N^2 + bx_N + c = y_N \end{array} \right\} \Rightarrow \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_N^2 & x_N & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \Rightarrow \mathbf{A}\mathbf{c} = \mathbf{y}$$

Once again, we solve the normal equation $\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{c} = \mathbf{A}^{\mathrm{T}}\mathbf{y}$ to get the least-squares approximate solution. This gives the system of equations:

$\begin{bmatrix} \sum_{i=1}^{N} x_i^4 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i & N \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} x_i^2 y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} y_i \end{bmatrix}$ which we then solve to find the coefficients a, b, c.

Example: Given the 5 data points $\{(1,1),(2,1),(3,1),(4,3),(5,5)\}$ find (a) the line that best fits this data and (b) the quadratic that best fits this data.

Solution: (a) It's easiest to assemble the necessary information in a table (or spreadsheet):

	x	y	x^2	xy
	1	1	1	1
	2	1	4	2
	3	1	9	3
	4	3	16	12
	5	5	25	25
Σ	15	11	55	43

If the line we seek has equation $y = mx + b$, the resulting normal equation is: $\begin{bmatrix} 55 & 15 \\ 15 & 5 \end{bmatrix}\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 43 \\ 11 \end{bmatrix}$.

We can easily solve this via row reduction or matrix inversion to get $m = 1$, $b = -.8$. So the line that best fits this data has equation $\boxed{y = x - .8}$.

(b) For the best-fitting quadratic we seek a parabola with equation $y = ax^2 + bx + c$. It's helpful to expand the previous table to get:

	x	y	x^2	xy	x^3	x^4	x^2y
	1	1	1	1	1	1	1
	2	1	4	2	8	16	4
	3	1	9	3	27	81	9
	4	3	16	12	64	256	48
	5	5	25	25	125	625	125
Σ	15	11	55	43	225	979	187

As previously described, the resulting normal equation becomes $\begin{bmatrix} 979 & 225 & 55 \\ 225 & 55 & 15 \\ 55 & 15 & 5 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 187 \\ 43 \\ 11 \end{bmatrix}$. Solving this with matrix inversion gives $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \dfrac{1}{70}\begin{bmatrix} 5 & -30 & 35 \\ -30 & 187 & -231 \\ 35 & -231 & 322 \end{bmatrix}\begin{bmatrix} 187 \\ 43 \\ 11 \end{bmatrix} = \dfrac{1}{70}\begin{bmatrix} 30 \\ -110 \\ 154 \end{bmatrix}$. So $a = \frac{3}{7}, b = -\frac{11}{7}, c = \frac{11}{5}$ and the best-fitting quadratic has equation $\boxed{y = \frac{3}{7}x^2 - \frac{11}{7}x + \frac{11}{5}}$.

More general least-squares methods

If a scatterplot suggests a relationship of the form $y = ax^p$ for some unknowns *a* and *p*, we can use logs to rewrite this as $\ln y = \ln a + p \ln x$. If we let $Y = \ln y$, $A = \ln a$, and $X = \ln x$, the relationship is then $Y = A + pX$ and we can use least-squares with the adjusted data to find *A* and *p*, and then exponentiate to find *a* and *p*.

These same methods work if we have data in the form $\{(x_i, y_i, z_i)\}_{i=1}^{N}$ and we're seeking the *plane* of best fit, or if we are trying to find the constants that provide a best fit for a relationship such as $z = ax^p y^q$ (in which case we would first take the log of both sides to get a relationship that yields a system of linear equations.

Notes by Robert Winters