INTRODUCTION TO GROUP REPRESENTATIONS
JUNE 18, 2012
LINEAR ALGEBRA REVIEW 2

The Standard Inner Product

Let $V = \mathbb{R}^n$ and $\langle \cdot, \cdot \rangle$ the standard inner product on $\mathbb{R}^n$. That is, if $v = (x_1, \ldots x_n)$ and $w = (y_1, \ldots y_n)$ are column vectors, then

$$\langle v, w \rangle = w^T v = \sum_i x_i y_i.$$

We state results for $\mathbb{R}^n$; many properties carry over directly to general inner product spaces. One verifies a routine list of properties, and, with the standard inner product, we can define lengths of vectors as

$$||v|| = \sqrt{\langle v, v \rangle} = \sum_i x_i^2$$

and the angle θ between vectors v and w in the plane spanned by v and w is obtained from

$$cos(\theta) = \frac{\langle v, w \rangle}{||v|| ||w||}.$$

In particular, v and w are called **orthogonal** if and only if $\langle v, w \rangle = 0$. In the plane, orthogonal vectors are perpendicular. If $||v|| = 1$, then v is called a **unit** vector.

As a first application, let $B = \{v_1, \ldots, v_n\}$ be a basis for $\mathbb{R}^n$. We call B an **orthonormal basis** for $\mathbb{R}^n$ if the v_i are mutually orthogonal and each v_i is a unit vector. As an example, consider the standard basis of $\mathbb{R}^n$.

If one has an orthonormal basis on hand (typically obtained from any basis using the Gram-Schmidt process), computations are somewhat simplified. We have

Fourier's Trick: if $v = c_1 v_1 + \cdots + c_n v_n$ then $c_i = \langle v, v_i \rangle$ and $||v||^2 = \sum_i c_i^2$.

Again one should verify in the case of the standard basis.

Orthogonal Complements

If W is a subspace of $\mathbb{R}^n$, then we define the **orthogonal complement** of W, denoted by $W^\perp$, as the subspace of all v in $\mathbb{R}^n$ such that

$$\langle v, w \rangle = 0 \quad \text{for all } w \ \text{ in } W.$$

Then we have an (orthogonal) direct sum decomposition

$$\mathbb{R}^n = W \oplus W^\perp;$$

Date: June 18, 2012.

that is, every v in $\mathbb{R}^n$ may be expressed uniquely as a sum $v_1 + v_2$ where v_1 is in W, v_2 is in $W^\perp$, and $\langle v_1, v_2\rangle = 0$. Furthermore if B is an orthonormal basis for W and B' an orthonormal basis for $W^\perp$, then $B \cup B'$ is an orthonormal basis for $\mathbb{R}^n$.

When $W = \mathbb{R}u$ for some nonzero unit vector u, then $v_1 = \langle v, u\rangle u$ and $v_2 = v - \langle v, u\rangle u$. This formula extends to general W: if we choose an orthonormal basis $B = \{u_1, \dots, u_k\}$ for W, then $v_1 = \sum_i \langle v, u_i\rangle u_i$ and $v_2 = v - \sum_i \langle v, u_i\rangle u_i$.

Orthogonal Projections

If A is an $n \times n$ matrix, we define the **adjoint** (or **transpose**) of A by $\langle Av, w\rangle = \langle v, A^T w\rangle$ for all v, w. If A has entries a_{ij}, then A^T has entries a_{ji}. If $A = A^T$, we call A **symmetric.**

For a fixed W, consider the map $P_W : \mathbb{R}^n \to W$ that sends $v = v_1 + v_2$ as above to v_1. We cal P_W the **orthogonal projection onto** W, and we note several properties of P_W:

(1) $P_W^2 = P_W$,
(2) $P_W + P_{W^\perp} = I$,
(3) $P_W P_{W^\perp} = P_{W^\perp} P_W = 0$, and
(4) $P_W^T = P_W$.

Property (4) is better understood through the inner product:

$$\langle P_W v, w\rangle = \langle P_W(v_1 + v_2), w_1 + w_2\rangle = \langle v_1, w_1 + w_2\rangle = \langle v_1, w_1\rangle = \langle v_1, P_W w\rangle = \langle v, P_W w\rangle.$$

To compute P_W as a matrix, we solve a diagonalization problem: if we choose orthonormal bases B for W and B' for $W^\perp$, then B is an orthonormal basis of eiegnvectors for eigenvalue 1 ($P_W v = 1 \cdot v$) and B' an orthonormal basis of eigenvectors for eigenvalue 0 ($P_W v = 0 \cdot v$). Suppose $B^* = B \cup B' = \{v_1, \dots, v_n\}$ and $P = [v_1 \dots v_n]$ is an orthogonal matrix. For an **orthogonal matrix** P, the following definitions are equivalent:

(1) $PP^T = P^T P = I$,
(2) $\langle Pv, Pw\rangle = \langle v, w\rangle$ for all v, w in $\mathbb{R}^n$,
(3) the columns of P form an orthonormal basis of $\mathbb{R}^n$, and
(4) the rows of P form an orthonormal basis of $\mathbb{R}^n$.

Now if $D = diag(1, \dots, 1, 0, \dots, 0)$ then

$$D = P^{-1} P_W P \quad \text{or} \quad P_W = PDP^{-1}.$$

Spectral Theorem

Consider the case of a diagonal transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(v) = Dv$ with D a real diagonal matrix. We can rewrite

$$T = \sum \lambda_i P_i,$$

where λ_i is the i-th diagonal entry and P_i is the orthogonal projection onto $\mathbb{R}e_i$. In other words,

$$D = \sum \lambda_i D_i,$$

where D_i is the zero everywhere save the (i, i)-th entry is 1.

There are two features of interest: we have taken a symmetric matrix (D) with real eigenvalues and diagonalized it with an orthonormal basis (the standard basis), and we have decomposed T into a sum of eigenvalues times orthogonal projections onto eigenspaces. This result holds in general for real symmetric matrices.

First observe what happens if we change to another orthonormal basis $C = \{w_1, \ldots, w_n\}$. Then $S = QDQ^{-1}$ with $Q = [w_1, \ldots w_n]$. Note that Q is orthogonal since C is orthonormal and that S is symmetric with eigenvectors $Qe_i = w_i$:

$$S^T = (QDQ^{-1})^T = (Q^{-1})^T D^T Q^T = QDQ^{-1} = S.$$

This process runs in reverse for any real symmetric matrix, in several steps:

(1) the eigenvalues of S are real,
(2) if W is invariant under S, then $W^\perp$ is also invariant under S,
(3) there exists at least one real eigenvalue for S, and
(4) we decompose $\mathbb{R}^n$ into one-dimensional subspaces invariant under S (eigenvectors for S) using induction.

When working over complex vector spaces, (3) requires no argument. Because we decompose by repeatedly forming orthogonal complements, the resulting basis B of eigenvectors is orthogonal with real eigenvalues and can be made orthonormal by rescaling. One also has that eigenvectors for distinct eigenvalues are orthogonal.

As in the motivating case, we decompose S along the eigenspaces. If P_i is the orthogonal projection onto the eigenspace for eigenvalue λ_i, then we have $S = \sum_i \lambda_i P_i$. If we apply the sum to an eigenvector v for eigenvalue λ_j, then

$$\sum_i \lambda_i P_i v = \lambda_j P_j v = \lambda_j v = Sv.$$

Orthogonal Matrix Decomposition

Fix an orthogonal matrix R. Consider matrices of the form

$$O = diag(I_j, -I_k, R(\theta_1), \ldots, R(\theta_l)),$$

where the $R(\theta)$ are orthogonal 2×2 blocks representing plane rotations. These are block diagonal, orthogonal, and represent the preferred similarity class for orthogonal matrices. That is, by adapting the proof of the Spectral Theorem, one may show that there exists another orthogonal matrix Q such that $O = Q^{-1}RQ$.

In the case of complex vector spaces, the result is much nicer if we replace "orthogonal" with the more general case of "unitary." Then O may be chosen to be diagonal with unit eigenvalues.

In the representation theory of finite groups, we generalize this result to full reducibility of representations. That is, if π is a unitary representation, we can decompose (π, V) into an orthogonal direct sum of irreducible representations (σ, V^σ) , the analogues of orthogonal eigenvector spans, such that

$$\pi(g) = \sum_{\sigma} \sigma(g) P_{\sigma}$$

holds for all g in G. With respect to an orthonormal basis, the unit eigenvalues are now replaced with unitary matrices for each $\sigma(g)$. Since every representation can be made unitary, the result applies in general.