An Introduction to Kostant's convexity theorem

Joseph Malkoun

NDU

April 7th, 2015

1. The Schur-Horn theorem
   - Hermitian matrices
   - The Schur-Horn theorem
2. Lie groups and Lie algebras
   - Lie groups
   - Lie algebras
   - Killing form
   - Semisimple Lie algebras
   - Cartan subalgebras
   - Maximal torus and Weyl group
3. Kostant's convexity theorem
4. Further generalizations

Hermitian matrices: their basic properties (1)

Definition

An hermitian n by n matrix A is a complex n by n matrix satisfying $A^* = A$, where $A^* = \bar{A}^T$.

Proposition

Eigenvalues of an hermitian matrix are real.

Proof.

Let $0 \neq v \in \mathbb{C}^n$ be an eigenvector of A with eigenvalue λ (A hermitian). Then

$$\bar{v}^T A v = \lambda \bar{v}^T v = \lambda \|v\|^2$$

Taking the conjugate transpose of the previous equation shows that

$$\bar{v}^T A^* v = \bar{\lambda} \|v\|^2 \Rightarrow \bar{\lambda} = \lambda$$

□

Hermitian matrices: their basic properties (2)

On $\mathbb{C}^n$, define the following map $< -, - >: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ by

$$< \mathbf{z}, \mathbf{w} >= \sum_{\alpha=1}^{n} z_\alpha \bar{w_\alpha}$$

$< -, - >$ is called the standard hermitian inner product on $\mathbb{C}^n$. It has the following properties:

1. $< c_1\mathbf{z}_1 + c_2\mathbf{z}_2, \mathbf{w} >= c_1 < \mathbf{z}_1, \mathbf{w} > + c_2 < \mathbf{z}_2, \mathbf{w} >$
2. $< \mathbf{z}, c_1\mathbf{w}_1 + c_2\mathbf{w}_2 >= \bar{c_1} < \mathbf{z}, \mathbf{w}_1 > + \bar{c_2} < \mathbf{z}, \mathbf{w}_2 >$
3. $< \mathbf{z}, \mathbf{w} >= \overline{< \mathbf{w}, \mathbf{z} >}$
4. $0 \neq \mathbf{z} \in \mathbb{C}^n \Rightarrow < \mathbf{z}, \mathbf{z} > > 0$

Hermitian matrices: their basic properties (3)

Proposition

Eigenvectors v_1, v_2 of an hermitian matrix A corresponding to different eigenvalues λ_1, λ_2, are orthogonal (i.e. $< v_1, v_2 >= 0$).

Proof.

$\bar{v_1}^T A v_2 = \lambda_2 < v_2, v_1 >$ and also $\bar{v_1}^T A v_2 = \bar{v_1}^T A^* v_2 = (\overline{Av_1})^T v_2 = \lambda_1 < v_2, v_1 >$. Hence $(\lambda_1 - \lambda_2) < v_2, v_1 >= 0$. But $\lambda_1 \neq \lambda_2$, so that $< v_2, v_1 >= 0$, and therefore $< v_1, v_2 >= \overline{< v_2, v_1 >} = 0$, as claimed. □

The finite-dimensional spectral theorem

Definition

An n by n matrix U is said to be unitary if $U^*U = \mathrm{Id}$.

Theorem

An hermitian n by n matrix A is unitarily diagonalizable. More precisely, there exist a diagonal matrix D and a unitary matrix U, such that $A = UDU^* = UDU^{-1}$. Then D consists of the n (real) eigenvalues $\lambda_1, \ldots, \lambda_n$ in some order in its diagonal entries, and U consists of corresponding eigenvectors.

Proof.

One can prove it by induction, using the fact that eigenvectors corresponding to different eigenvalues $\lambda \neq \lambda'$ are orthogonal, with respect to the standard hermitian inner product on $\mathbb{C}^n$, and a Gram-Schmidt orthogonalization in each eigenspace of dimension greater than 1. □

The Schur-Horn theorem (1)

Given an n by n hermitian matrix A, one can compute its eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$, which are determined up to a permutation, and one can also extract the diagonal vector $\mathbf{d} = (a_{11}, \ldots, a_{nn}) \in \mathbb{R}^n$. The Schur-Horn theorem provides necessary and sufficient conditions for two vectors λ and $\mathbf{d}$ in $\mathbb{R}^n$ to be respectively the eigenvalues and diagonal for some hermitian matrix.

The Schur-Horn theorem (2)

Theorem (Schur-Horn theorem)

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\mathbf{d} = (d_1, \ldots, d_n)$ be two points in $\mathbb{R}^n$. Reorder the λ_i, so that the reordered λ_i, which we denote by a prime, satisfy $\lambda'_1 \geq \lambda'_2 \geq \ldots \geq \lambda'_n$. Do the same for the d_i, so that $d'_1 \geq d'_2 \geq \ldots \geq d'_n$. Then there exists an hermitian n by n matrix A having as eigenvalues λ and as diagonal $\mathbf{d}$ iff

$$d'_1 \leq \lambda'_1 \tag{2.1}$$
$$d'_1 + d'_2 \leq \lambda'_1 + \lambda'_2 \tag{2.2}$$
$$\vdots \tag{2.3}$$
$$d'_1 + \cdots + d'_{n-1} \leq \lambda'_1 + \cdots + \lambda'_{n-1} \tag{2.4}$$
$$d'_1 + \cdots + d'_n = \lambda'_1 + \cdots + \lambda'_n \tag{2.5}$$

The Schur-Horn theorem (3)

A subset $C \subseteq \mathbb{R}^n$ is said to be convex if given any two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in C, the points $(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in C$ for all t such that $0 \leq t \leq 1$. In other words, the line segment joining any two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in C must be entirely contained in C. Given a subset $S \subseteq \mathbb{R}^n$, we define its convex hull $\hat{S}$ as the smallest convex subset of $\mathbb{R}^n$ containing S. A convex polytope is the convex hull of a finite set S. As geometer D. Coxeter used to say, the term polytope is the general term of the sequence: "point", "polygon", "polyhedron", etc.

The Schur-Horn theorem (4)

Let Σ_n be the group of all permutations of the set $\{1, \ldots, n\}$. Σ_n acts on $\mathbb{R}^n$ by permuting the n standard coordinates $x_1, \ldots, x_n$ of $\mathbb{R}^n$. Given two points λ and $\mathbf{d}$ in $\mathbb{R}^n$, as in the Schur-Horn theorem, it turns out that the geometric conditions in that theorem are equivalent to d being in the convex hull of the orbit $\Sigma_n.\lambda$, the latter convex hull being a convex polytope, since the orbit $\Sigma_n.\lambda$ is finite.

The Schur-Horn theorem (5)

We can now reformulate the Schur-Horn theorem in more geometric terms.

Theorem (Schur-Horn, reformulated)

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\mathbf{d} = (d_1, \ldots, d_n)$ be two points in $\mathbb{R}^n$. Then there exists an hermitian n by n matrix A having as eigenvalues λ and as diagonal $\mathbf{d}$ iff $\mathbf{d}$ is in the convex hull $\hat{S}$ of S, where

$$S = \{(\lambda_{\sigma^{-1}(1)}, \ldots, \lambda_{\sigma^{-1}(n)}); \sigma \in \Sigma_n\}$$

Basic notions (1)

Definition (manifold)

A smooth n-manifold M is a second countable Hausdorff space, covered by local charts $\{(U, f_U)\}$, where $U \subseteq M$ is open and $f_U : U \to U'$, $U' \subseteq \mathbb{R}^n$ open, and f_U homeomorphism, such that the transition function $f_V \circ f_U^{-1}$, restricted to $f_U(U \cap V)$ is a C^∞ diffeomorphism from $f_U(U \cap V)$ onto $f_V(U \cap V)$, whenever $U \cap V \neq \emptyset$.

From now on, we always assume that our manifolds are finite-dimensional. In particular, Lie groups will be finite-dimensional, etc.

Basic notions (2)

Definition (Lie group)

A Lie group G is a group which is also a smooth manifold, such that both multiplication $G \times G \to G$ and the map $G \to G$ sending g to g^{-1} are both C^∞.

Examples of Lie groups (1)

Example (orthogonal groups)

The real orthogonal group $O(n, \mathbb{R})$ is defined by

$$O(n) = \{g \in GL(n, \mathbb{R}); g^T g = \mathrm{Id}\}$$

If $(-, -)$ denotes the standard inner product on $\mathbb{R}^n$, then the orthogonal group is the group of all linear automorphisms g of $\mathbb{R}^n$ which preserve $(-, -)$, i.e. such that $(g(\mathbf{v}), g(\mathbf{w})) = (\mathbf{v}, \mathbf{w})$ for any $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. If $g \in O(n, \mathbb{R})$, then $\det(g) = \pm 1$ (by taking the determinant on both sides of $g^T g = \mathrm{Id}$). The group $O(n)$ has two connected components. The special orthogonal group $SO(n)$ is $SO(n) = \{g \in O(n); \det(g) = 1\}$. The Lie group $SO(n)$ is a compact *connected* Lie group of (real) dimension $n(n-1)/2$ (while $O(n)$ is not connected).

Examples of Lie groups (2)

Example (unitary groups)

The unitary group $U(n)$ is defined by

$$U(n) = \{g \in GL(n, \mathbb{C}); g^* g = \mathrm{Id}\}$$

Using the standard hermitian inner product $< -, - >$ defined earlier, the group $U(n)$ is the group of $\mathbb{C}$-linear automorphisms g of $\mathbb{C}^n$ which preserve $< -, - >$, i.e. $< g(\mathbf{v}), g(\mathbf{w}) >=< \mathbf{v}, \mathbf{w} >$. By taking the determinant on both sides of the defining equation, we get that $|\det(g)| = 1$ for all $g \in U(n)$. We define also define $SU(n) = \{g \in U(n); \det(g) = 1\}$. $U(n)$ and $SU(n)$ are compact connected Lie groups of real dimensions n^2 and $n^2 - 1$ respectively.

Lie algebras

Definition

A (real) Lie algebra is a pair $(V, [-,-])$ where V is a real vector space, and $[-,-] : V \times V \to V$ is a map satisfying:

1. (bilinearity) for any fixed $x \in V$, $[x,-] : V \to V$ is $\mathbb{R}$-linear, and similarly $[-,x] : V \to V$ is $\mathbb{R}$-linear,
2. (skew-symmetry) for any $x, y \in V$, $[x,y] = -[y,x]$,
3. (Jacobi identity) for any $x, y, z \in V$, $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$.

The Lie algebra associated to a Lie group (1)

Proposition

One can associate to any Lie group G a Lie algebra $\mathfrak{g}$ of the same dimension.

If G is a Lie group, denote by $\mathfrak{g} = T_1(G)$, where 1 is the identity element of G. We wish to define a Lie bracket $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. First we define the Adjoint action of G on itself. For every $g \in G$, define $Ad_g : G \to G$ by $Ad_g(g_1) = gg_1g^{-1}$. Then Ad is a smooth group action of G on itself. It is also clear that $Ad_g(1) = 1$ for any $g \in G$. For a given $g \in G$, one can differentiate Ad_g at 1, this gives an $\mathbb{R}$-linear map $(Ad_g)_* : T_1(G) \to T_1(G)$, in other words $(Ad_g)_* : \mathfrak{g} \to \mathfrak{g}$. This is the adjoint action of G on $\mathfrak{g}$.

The Lie algebra associated to a Lie group (2)

We now differentiate again $(Ad_g)_*$ at $g = 1$. This gives an $\mathbb{R}$-linear map $ad : \mathfrak{g} \to \text{End}_{\mathbb{R}}(\mathfrak{g})$ called the adjoint action of $\mathfrak{g}$ on itself. We now define $[x, y] = ad_x(y)$ for $x, y \in \mathfrak{g}$. One can check that this bracket is actually a Lie bracket. $(\mathfrak{g}, [-, -])$ is the Lie algebra associated to G.

The Lie algebra associated to a Lie group (3)

From the previous proof, we see that a Lie group G acts on its Lie algebra $\mathfrak{g}$, where an element $g \in G$ acts via $(Ad_g)_* \in \mathrm{End}_{\mathbb{R}}(\mathfrak{g})$. Actually, for any $g \in G$, Ad_g is invertible, with inverse $Ad_{g^{-1}}$. It is customary to drop the star from the notation, and simply write Ad_g, for the action of g on the Lie algebra $\mathfrak{g}$.

The Adjoint action, more concretely (1)

Now is a good time to look at matrix groups, to make things more concrete. By a (real) matrix group, we mean a subgroup of $GL(n,\mathbb{R})$, the general linear group consisting of real linear automorphisms of $\mathbb{R}^n$. $GL(n,\mathbb{R})$ is itself a Lie group of dimension n^2. If $G = GL(n,\mathbb{R})$, then for $g, g_1 \in G$, we have $Ad_g(g_1) = gg_1g^{-1}$, using matrix multiplication. Writing $g_1 = \mathrm{Id} + ty$, where $y \in \mathfrak{g} = \mathfrak{gl}(n,\mathbb{R})$, the space of *all* n by n real matrices, we get $Ad_g(g_1) = Ad_g(\mathrm{Id} + ty) = \mathrm{Id} + tAd_g(y) = \mathrm{Id} + tgyg^{-1}$, so that the Adjoint action of G on $\mathfrak{g}$ is given by $Ad_g(y) = gyg^{-1}$. Replacing g by $\mathrm{Id} + \tau x$, where $x \in \mathfrak{g}$, we see that

$$\begin{aligned} Ad_{\mathrm{Id}+\tau x}(y) &= (\mathrm{Id} + \tau x)y(\mathrm{Id} + \tau x)^{-1} \\ &= (\mathrm{Id} + \tau x)y(\mathrm{Id} - \tau x + O(\tau^2)) \\ &= y + \tau(xy - yx) + O(\tau^2) \end{aligned}$$

The Adjoint action, more concretely (2)

Summarizing, if $G = GL(n, \mathbb{R})$, then the Adjoint action of G on its Lie algebra $\mathfrak{g}$ is given by $Ad_g(y) = gyg^{-1}$, where $g \in G$ and $y \in \mathfrak{g}$, and the adjoint action of $\mathfrak{g}$ on itself is given by $ad_x(y) = [x, y] = xy - yx$. We shall be interested in the Adjoint action of G on the dual $\mathfrak{g}^*$ of $\mathfrak{g}$. An element $g \in G$ maps $\xi \in \mathfrak{g}^*$ to $Ad^*_{g^{-1}}(\xi) \in \mathfrak{g}^*$, which satisfies $(x, Ad^*_{g^{-1}}(\xi)) = (Ad_{g^{-1}}(x), \xi) = (g^{-1}xg, \xi)$, for all $x \in \mathfrak{g}$, where $(-, -)$ here denotes the natural pairing $\mathfrak{g} \times \mathfrak{g}^* \to \mathbb{R}$ (which is a non-degenerate bilinear pairing).

Some examples of Lie algebras (1)

We have previously defined the special orthorgonal groups $SO(n)$. We now wish to work out its associated Lie algebra $\mathfrak{so}(n)$. We have two equations to differentiate, $g^T g = \mathrm{Id}$ and $\det(g) = 1$. Replace g by $\mathrm{Id} + tx$, where x is a real n by n matrix, and t is a small real parameter. Then $(\mathrm{Id} + tx)^T(\mathrm{Id} + tx) = \mathrm{Id}$, so $\mathrm{Id} + t(x^T + x) + O(t^2) = \mathrm{Id}$, from which we deduce that $x + x^T = 0$. In other words, x is skew-symmetric. Let us consider now the remaining equation $\det(g) = 1$. This gives $\det(\mathrm{Id} + tx) = 1$, so $1 + t\,\mathrm{tr}(x) + O(t^2) = 1$, from which we deduce that $\mathrm{tr}(x) = 0$. But this was already implied by the equation $x + x^T = 0$ by taking the trace on both sides.

Some examples of Lie algebras (2)

We have thus found that

$$\mathfrak{so}(n) = \{x \in \mathfrak{gl}(n, \mathbb{R}); x + x^T = 0\}$$

We can now easily see that the dimension of $\mathfrak{so}(n)$ is $n(n-1)/2$, which is also the dimension of $SO(n)$. Being a subgroup of $GL(n, \mathbb{R})$, it follows that the Lie bracket of $\mathfrak{so}(n)$ is also the commutator $[x, y] = xy - yx$, for $x, y \in \mathfrak{so}(n)$.

Some examples of Lie algebras (3)

In a similar way, we find that

$$\mathfrak{u}(n) = \{x \in \mathfrak{gl}(n, \mathbb{C}); x + x^* = 0\}$$

Thus $\mathfrak{u}(n)$ consists of skew-hermitian complex n by n matrices. We also have

$$\mathfrak{su}(n) = \{x \in \mathfrak{u}(n); \mathrm{tr}(x) = 0\}$$

The Lie brackets for both $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$ are given by the commutator $[x, y] = xy - yx$. We can now compute the real dimension of $\mathfrak{u}(n)$ to be n^2, and that of $\mathfrak{su}(n)$ to be $n^2 - 1$. We also remark that multiplication by i identifies the space of hermitian n by n matrices with the space of n by n skew-hermitian matrices, the latter being $\mathfrak{u}(n)$. This already provides a clue as to how Kostant generalized the Schur-Horn theorem!

The Killing form of a Lie algebra

Given a (real) Lie algebra $(\mathfrak{g}, [-,-])$, for each element $x \in \mathfrak{g}$, $ad_x \in \mathrm{End}(\mathfrak{g})$. The Killing form of the Lie algebra $\mathfrak{g}$ is the map $(-,-) : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ defined by

$$(x, y) = \mathrm{tr}(ad_x \circ ad_y)$$

The Killing form $(-,-)$ is a symmetric bilinear form on $\mathfrak{g}$.

Some examples of Killing forms (1)

Consider the Lie algebra $\mathfrak{so}(n)$. Using the definition and a small computation, one can show that its Killing form is

$$(x, y) = (n-2)\,\mathrm{tr}(xy)$$

We remark that $(x, x) = (n-2)\,\mathrm{tr}(x^2) = -(n-2)\,\mathrm{tr}(x^T x) < 0$ if $x \neq 0$, so that the Killing form of $\mathfrak{so}(n)$ is negative-definite.

Some examples of Killing forms (2)

Consider the Lie algebra $\mathfrak{u}(n)$. Its Killing form can be shown to be

$$(x, y) = 2(n\,\mathrm{tr}(xy) - \mathrm{tr}(x)\,\mathrm{tr}(y))$$

which is a real-valued negative semi-definite bilinear form, but degenerate. Indeed, $i\,\mathrm{Id} \in \mathfrak{u}(n)$, and $(x, i\,\mathrm{Id}) = 0$ for all $x \in \mathfrak{u}(n)$. On the other hand, the Killing form of $\mathfrak{su}(n)$ is

$$(x, y) = 2n\,\mathrm{tr}(xy)$$

which is negative-definite.

Semisimple Lie algebras (1)

Definition

A subset $\mathfrak{I} \subseteq \mathfrak{g}$ is said to be an ideal of $\mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{I}] \subseteq \mathfrak{I}$. A Lie algebra $\mathfrak{g}$ is said to be simple if it is non-abelian (its Lie bracket does not vanish identically) and if its only ideals are $\mathfrak{g}$ and $\mathbf{0}$. If $\mathfrak{g}_1$, $\mathfrak{g}_2$ are two Lie algebras, then their direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ is also a Lie algebra, with Lie bracket defined by $[(x_1, x_2), (y_1, y_2)] = ([x_1, y_1]_{\mathfrak{g}_1}, [x_2, y_2]_{\mathfrak{g}_2})$. A Lie algebra $\mathfrak{g}$ is said to be semisimple if it is the direct sum of simple Lie algebras.

Semisimple Lie algebras (2)

The Lie algebras $\mathfrak{su}(n)$ $(n \geq 2)$ and $\mathfrak{so}(n)$ $(n \geq 3)$ are semisimple. The Lie algebras $\mathfrak{gl}(n)$ and $\mathfrak{u}(n)$ are not semisimple, since they have a non-trivial center (the center of $\mathfrak{g}$ consists of all elements $x \in \mathfrak{g}$ such that $[x, \mathfrak{g}] = 0$). It is interesting to note that while $\mathfrak{so}(n)$ is simple for $n = 3$ and for $n \geq 5$, but for $n = 4$, something interesting happens:

$$\mathfrak{so}(4) = \mathfrak{so}(3) \oplus \mathfrak{so}(3)$$

Cartan's criterion

Theorem (Cartan's criterion)

A Lie algebra is semisimple iff its Killing form is non-degenerate.

We have seen that the Killing forms of $\mathfrak{so}(n)$ and $\mathfrak{su}(\mathfrak{n})$ are negative-definite, and therefore non-degenerate, so that $\mathfrak{so}(n)$ and $\mathfrak{su}(n)$ are semisimple, as previously claimed.

Cartan subalgebras

We now return to the general setting: let $\mathfrak{g}$ be a Lie algebra (not necessarily semisimple).

Definition

$\mathfrak{g}$ is said to be nilpotent if the sequence defined by $\mathfrak{g}_0 = \mathfrak{g}$, $\mathfrak{g}_n = [\mathfrak{g}, \mathfrak{g}_{n-1}]$ is $\mathbf{0}$ after finitely many n. A subspace $\mathfrak{h} \subseteq \mathfrak{g}$ is said to be a subalgebra if it is closed under the Lie bracket of $\mathfrak{g}$. A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is said to be self-normalizing if whenever $x \in \mathfrak{g}$ satisfies $[x, \mathfrak{h}] \subseteq \mathfrak{h}$, then $x \in \mathfrak{h}$. A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is said to be a Cartan subalgebra if it is nilpotent and self-normalizing.

Maximal torus of a compact Lie group

Let G be a compact Lie group.

Definition

A torus $T \subseteq G$ is a compact connected abelian Lie subgroup of G. A torus T in G is said to be maximal if it is maximal in the sense of inclusion (so that there does not exist a torus T' containing T other than T itself).

Given a torus T in G, the Weyl group $W(T, G)$ is defined to be

$$W(T, G) = N(T)/C(T)$$

where $N(T)$ is the normalizer of T in G, and $Z(T)$ is the centralizer of T in G. It turns out that if G is a compact connected Lie group, then any two maximal tori are conjugate by some element in G.

Examples of maximal tori and Weyl groups (1)

- If $G = U(n)$, then the diagonal matrices in $U(n)$, namely matrices of the form
$$\begin{pmatrix} e^{i\theta_1} & 0 & \cdots & 0 \\ 0 & e^{i\theta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{i\theta_n} \end{pmatrix}$$
form a maximal torus T in $U(n)$. Then the Weyl group $W(T, G) \simeq \Sigma_n$, the permutation group on n elements.
- If $G = SU(n)$, then the diagonal matrices in that group also form a maximal torus T. They consist of matrices of the same form, but also having determinant 1. We also have $W(T, G) \simeq \Sigma_n$.

Examples of maximal tori and Weyl groups (2)

- If $G = SO(2n)$, then a maximal torus T is given by matrices of the form
$$\begin{pmatrix} a_1 & -b_1 & \cdots & 0 & 0 \\ b_1 & a_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_n & -b_n \\ 0 & 0 & \cdots & b_n & a_n \end{pmatrix}$$
where $a_i^2 + b_i^2 = 1$, $1 \leq i \leq n$. In this case, the Weyl group $W(T, G) \simeq \Sigma_n \ltimes (\mathbb{Z}_2)^{n-1}$, with $\sigma s_i s_j \sigma^{-1} = s_{\sigma^{-1}(i)} s_{\sigma^{-1}(j)}$, for $\sigma \in \Sigma_n$ and $s_i s_j \in (\mathbb{Z}_2)^{n-1}$ $(i \neq j)$ has all ones except at i and j.

Examples of maximal tori and Weyl groups (3)

- If $G = SO(2n+1)$, then a maximal torus T is given by matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_1 & -b_1 & \cdots & 0 & 0 \\ 0 & b_1 & a_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_n & -b_n \\ 0 & 0 & 0 & \cdots & b_n & a_n \end{pmatrix}$$

where $a_i^2 + b_i^2 = 1$, $1 \leq i \leq n$. In this case, the Weyl group $W(T, G) \simeq \Sigma_n \ltimes (\mathbb{Z}_2)^n$, with $\sigma s_i \sigma^{-1} = s_{\sigma^{-1}(i)}$, for $\sigma \in \Sigma_n$ and $s_i \in (\mathbb{Z}_2)^n$ $(1 \leq i \leq n)$ has all ones except at i.

Kostant's convexity theorem (1)

This section draws material from an article by Francois Ziegler in the Proceedings of the AMS, entitled "On the Kostant convexity theorem" (1992). Let G be a compact connected Lie group, T a maximal torus of G, with associated Lie algebras $\mathfrak{g}$ and $\mathfrak{t}$. Let $\pi : \mathfrak{g}^* \to \mathfrak{t}^*$ be the natural projection. Then $\mathfrak{t}^*$ can be identified with the subspace of all T-fixed points in $\mathfrak{g}^*$. Every coadjoint X of G intersects $\mathfrak{t}^*$ in a Weyl group orbit Ω_X. B. Kostant has proved that

Theorem (Kostant convexity theorem)

$\pi(X)$ is the convex hull of Ω_X.

Kostant's convexity theorem (2)

Let us apply the theorem to $G = U(n)$ and T consisting of all unitary n by n diagonal matrices. The real-valued symmetric bilinear form $(-,-)$ on $\mathfrak{g}$ mapping $(x, y) = \mathrm{tr}(x^* y)$ is positive definite and Ad_G-invariant. It allows us to identify $\mathfrak{g} \simeq \mathfrak{g}^*$. We have already seen that by multiplying by i, the space of hermitian n by n matrices can be identified with $\mathfrak{g}$, the latter being the space of n by n skew-hermitian matrices.

Kostant's convexity theorem (3)

Coadjoint orbits correspond to isospectral sets of hermitian matrices, which are hermitian matrices having the same fixed eigenvalues $\lambda_1, \cdots, \lambda_n$. In this case, the image of a coadjoint orbit under π just consists of the diagonals of all the matrices in an isospectral set of hermitian matrices. On the other hand, the Weyl orbit Ω_X just consists of $\Sigma_n.(\lambda_1, \cdots, \lambda_n)$. This shows that Kostant's convexity theorem is indeed a Lie-theoretic generalization of the Schur-Horn theorem.

Further generalizations

Atiyah and independently Guillemin and Sternberg proved, almost simultaneously in 1982, a generalization of Kostant's convexity theorem in the setting of compact symplectic manifolds having a hamiltonian toric action. The image of such a manifold under the moment map is then also a convex polytope (more precisely, it is the convex hull of the images of the fixed points of the manifold under the torus action).

And finally...

Thank you!!!