II. Linear Algebraic Groups

You may be familiar with the notion of a *Lie group*, which is a differentiable manifold that also admits a group structure. Most of the interesting Lie groups, and all of the *simple* ones, admit matrix representations. In the classification of simple complex Lie groups by Cartan and Killing there are four infinite series A_n, B_n, C_n and D_n of *classical groups* and the five *exceptional groups* G_2, F_4, E_6, E_7, and E_8. These are the symmetry groups of "continuous" phenomena in mathematics and physics.

However, one sometimes needs to study arithmetic phenomena, or work algebraically. In that case the realization of these groups as matrix groups allows us to restrict the coefficients (i.e. matrix entries) to lie in a particular field — not just subfields of $\mathbb{C}$.[1] We will take this more general point of view and restrict our study to groups of matrices preserving a (nondegenerate) symmetric or alternating bilinear form, in parallel with the complex classical Lie groups.

Where things become more complicated in the classical theory is when one gets into *real forms* of the complex Lie groups. The point is that there are distinct real Lie groups whose "complexifications" yield the same complex Lie group, starting with $\mathbb{R}^*$ and the unit circle S^1 (which are both real forms of $\mathbb{C}^*$). Unitary groups, which we will touch on, are real forms of special linear groups. Over more general fields, these "forms" of linear algebraic groups proliferate even further, and interact with Galois theory, even if we shall largely avoid this in our brief study.

[1] One can abstractify much further and define *algebraic groups* to be "schemes" whose $\mathbb{F}$-points form a group for any field $\mathbb{F}$ extending their field of definition. (Obviously, we aren't going to do that here.) We could also ask for entries in $\mathbb{Z}$ or some $\mathcal{O}_K$ to get discrete "arithmetic groups".

II.A. Bilinear forms

Let V be a finite-dimensional vector space over a field $\mathbb{F}$, with basis $e = \{e_1, \dots, e_n\}$. Its **dual space** is the vector space

$$V^\vee := \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$$

of $\mathbb{F}$-linear functionals. This has a (dual) basis $e^\vee = \{e_1^\vee, \dots, e_n^\vee\}$ defined by $e_i^\vee(e_j) = \delta_{ij}$; for an arbitrary $v = \sum_j v_j e_j \in V$, this yields $e_i^\vee(v) = v_i$.

II.A.1. NOTATION. We will denote coordinate vectors (of elements of V or $V^\vee$) with respect to a basis b by $[\,\cdot\,]_b$. With $f = \sum_i f_i e_i^\vee \in V^\vee$, and v as above, this gives:

- $[v]_e = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$;
- ${}^t[f]_{e^\vee} = (f_1, \dots, f_n)$; and thus (by matrix multiplication)
- ${}^t[f]_{e^\vee}[v]_e = f(v)$.

Similarly, matrices of a transformation, say $T \in \mathrm{Hom}_{\mathbb{F}}(W, W')$, with respect to bases b of W and b' of W', are written ${}_{b'}[T]_b$. If T has an inverse, then ${}_b[T^{-1}]_{b'} = ({}_{b'}[T]_b)^{-1}$.

Here we'll mainly use this as follows: given a second basis $\tilde{e}$ of V, and writing $\mathbf{1} \in \mathrm{Hom}_{\mathbb{F}}(V, V)$ for the identity map, the *change-of-basis matrix* ${}_{\tilde{e}}[\mathbf{1}]_e$ satisfies ${}_{\tilde{e}}[\mathbf{1}]_e[v]_e = [v]_{\tilde{e}}$. (We will also write $\mathbf{1}$ for the identity map on $V^\vee$.) Computing $f(v)$ two ways

- $f(v) \;=\; {}^t[f]_{\tilde{e}^\vee}[v]_{\tilde{e}} \;=\; {}^t[f]_{\tilde{e}^\vee}\,{}_{\tilde{e}}[\mathbf{1}]_e\,[v]_e$
- $f(v) \;=\; {}^t[f]_{e^\vee}[v]_e \;=\; {}^t({}_{e^\vee}[\mathbf{1}]_{\tilde{e}^\vee}\,[f]_{\tilde{e}^\vee})[v]_e \;=\; {}^t[f]_{\tilde{e}^\vee}\,{}^t_{e^\vee}[\mathbf{1}]_{\tilde{e}^\vee}\,[v]_e$

yields (since it holds for all $f \in V^\vee$ and $v \in V$) the formula

$${}_{e^\vee}[\mathbf{1}]_{\tilde{e}^\vee} \;=\; {}^t_{\tilde{e}}[\mathbf{1}]_e \;=\; {}^t({}_{\tilde{e}}[\mathbf{1}]_e)^{-1}.$$

II.A.2. DEFINITION. (i) The **tensor product** of V with itself, written $V \otimes V$, is the $\mathbb{F}$-vector space with basis $\{e_i \otimes e_j\}_{i,j=1,\dots,n}$. So as a set

$$V \otimes V = \{\textstyle\sum_{i,j} a_{ij} e_i \otimes e_j \mid a_{ij} \in \mathbb{F}\}.$$

Given $v = \sum_i v_i e_i$ and $w = \sum_j w_j e_j$ in V, we can define an element[2] $v \otimes w \in V \otimes V$ by

$$v \otimes w = (\textstyle\sum_i v_i e_i) \otimes (\sum_j w_j e_j) := \sum_{i,j} v_i w_j \, e_i \otimes e_j.$$

When there is more than one field over which one might consider V as a vector space, we write $V \otimes_{\mathbb{F}} V$ to disambiguate the notation.

(ii) A **bilinear form** V is an element

$$B \in \mathrm{Hom}_{\mathbb{F}}(V \otimes V, \mathbb{F}) \ (\cong V^\vee \otimes V^\vee).$$

This is equivalent to giving a map $B(\cdot,\cdot)\colon V \times V \to \mathbb{F}$ which is separately $\mathbb{F}$-linear in each entry (why?). The *matrix* of B with respect to e is simply $[B]_e := (B(e_i, e_j))$, and

$$B(v,w) = B(\textstyle\sum_i v_i e_i, \sum_j w_j e_j) = \sum_{i,j} v_i w_j \, B(e_i, e_j) = {}^t[v]_e \, [B]_e \, [w]_e.$$

Under change of basis from e to $\tilde{e}$, and writing $S := {}_e[\mathbf{1}]_{\tilde{e}}$, we have

$${}^t[v]_{\tilde{e}}[B]_{\tilde{e}}[w]_{\tilde{e}} = B(v,w) = {}^t[v]_e[B]_e[w]_e = {}^t[v]_{\tilde{e}} \, {}^tS[B]_e S[w]_{\tilde{e}}$$

for every $v, w \in V$ hence $[B]_{\tilde{e}} = {}^tS[B]_e S$.

(iii) The **discriminant** $\Delta_B := \det[B]_e$ of the bilinear form is only well-defined up to multiplication by a square, since the change-of-basis formula yields $\det[B]_{\tilde{e}} = (\det S)^2 \det[B]_e$ for some matrix S (and we want something invariant with respect to choice of basis). So we regard Δ_B as an element of $\mathbb{F}^* / (\mathbb{F}^*)^2 \cup \{0\}$.

A bilinear form determines homomorphisms

$$B_L\colon V \to V^\vee \qquad \text{and} \qquad B_R\colon V \to V^\vee$$
$$v \mapsto B_L(v) := B(v,\cdot) \qquad\qquad v \mapsto B_R(v) := B(\cdot, v)$$

which are easy to describe in matrix terms. For instance, since

$${}^t[B_R(v)]_{e^\vee}[w]_e = B_R(v)(w) = B(w,v) = {}^t[w]_e[B]_e[v]_e = {}^t[v]_e \, {}^t[B]_e[w]_e,$$

[2]Warning: not all elements of $V \otimes V$ are of this form. They are like matrices of rank 1.

we get $[B_R(v)]_{e^\vee} = [B]_e[v]_e$, hence ${}_{e^\vee}[B_R]_e = [B]_e$; and similarly, one has ${}_{e^\vee}[B_L]_e = {}^t[B]_e$. At this level of generality, "orthogonal" subspaces are pretty weird, but one can define them: if $U \subseteq V$ is any $\mathbb{F}$-subspace, then

$$U^{\perp_L} := \bigcap_{u \in U} \ker(B_L(u)) \quad \text{and} \quad U^{\perp_R} := \bigcap_{u \in U} \ker(B_R(u))$$

are subspaces of V. One nice property is that, since U "is in the right kernel of everything in its left kernel", we have $U \subseteq (U^{\perp_L})^{\perp_R}$ (and vice-versa). The subspaces $V^{\perp_L}$ and $V^{\perp_R}$, in particular, are called the **left** and **right radical** of V.

II.A.3. PROPOSITION-DEFINITION. *The following are equivalent:*
(i) $V^{\perp_L} = \{0\}$*;*
(ii) $V^{\perp_R} = \{0\}$*; and*
(iii) $\Delta_B \neq 0$ ($[B]_e$ *invertible*).
When any (*hence all of them*) *hold, B is said to be* ***nondegenerate****.*

PROOF. Clearly (i) holds $\iff$ $\cap_{v \in V} \ker(B_L(v)) = \{0\}$. This translates to the statement that no coordinate vector $[w]_e$ has zero dot product with every coordinate vector ${}^t[B]_e[v]_e$, i.e. with every vector in the (column space of ${}^t[B]$ =) row space of $[B]_e$. This is the same as saying the nullspace of $[B]_e$ is zero, i.e. that $[B]_e$ is invertible. The equivalency (ii) $\iff$ (iii) is similar. □

II.A.4. PROPOSITION. *For B nondegenerate, and $U \subset V$ a subspace:*
(a) *Any $f \in \mathrm{Hom}_{\mathbb{F}}(U, \mathbb{F}) = U^\vee$ is the restriction to U of $B_L(w)$ for some $w \in V$* (*and $B_R(w')$ for some $w' \in V$*)*.*
(b) $(\cdot)^{\perp_R}$ *and* $(\cdot)^{\perp_L}$ *are mutual inverses* (*hence both bijective*) *on the set of subspaces of V.*

PROOF. Let $\tilde{e}_1, \ldots, \tilde{e}_\ell$ be a basis of U, and $\tilde{e}_{\ell+1}, \ldots, \tilde{e}_n$ the rest of a basis for V. The choice of those remaining basis elements gives a choice of extension of $\tilde{e}_i^\vee \in U^\vee$ to $\tilde{e}_i^\vee \in V^\vee$, hence for any $f = \sum_{i=1}^{\ell} f_i \tilde{e}_i^\vee \in U^\vee$ to an element of $V^\vee$. In particular, the restriction-of-functionals map $\imath^\vee \colon V^\vee \to U^\vee$ dual to $\imath \colon U \hookrightarrow V$ is surjective. So by "rank + nullity", $\dim(\ker \imath^\vee) = n - \ell$.

Moreover, $U^{\perp_R}$ is the kernel of the composition

$$\begin{array}{ccccc} V & \xrightarrow[\cong]{B_R} & V^\vee & \twoheadrightarrow & U^\vee \\ v & \longmapsto & B(\,\cdot\,, v) & & \end{array}$$

so has dimension $n - \ell$. (This holds for any subspace and also replacing R by L). But then $(U^{\perp_R})^{\perp_L} \supseteq U$ have the same dimension ($n - (n - \ell) = \ell$) so are equal. □

II.A.5. DEFINITION. Here are three (essentially two) special kinds of bilinear forms:

- B is **symmetric** $\iff B(x,y) = B(y,x)$ $(\forall x, y \in V)$.
- B is **alternating** $\iff B(x,x) = 0$ $(\forall x \in V)$.
- B is **skew-symmetric** $\iff B(x,y) = -B(y,x)$ $(\forall x, y \in V)$.

That alternating implies skew-symmetric follows from

$$B(x,y) + B(y,x) = B(x+y, x+y) - B(x,x) - B(y,y).$$

If $\mathrm{char}(\mathbb{F}) \neq 2$, then skew-symmetric forms are alternating (put $x = y$, and divide by 2). Note that B is symmetric [resp. skew-symmetric] if and only if its matrix with respect to any basis satisfies ${}^t[B]_e = [B]_e$ [resp. $-[B]_e$].

Given a bilinear form B on V, "$x \perp y$" will mean $B(x,y) = 0$.

II.A.6. THEOREM. *The following are equivalent:*
(i) $x \perp y \iff y \perp x$ *for all* $x, y \in V$.
(ii) B *is symmetric or alternating.*

PROOF. Obviously (ii) $\implies$ (i). So assume (i), let $x, y, z \in V$ be arbitrary, and put $v := B(x,y)z - B(x,z)y$. Then

$$B(x,v) = B(x,y)B(x,z) - B(x,z)B(x,y) = 0 \implies x \perp v$$
$$\implies v \perp x \implies 0 = B(v,x) = B(x,y)B(z,x) - B(x,z)B(y,x),$$

and taking $y := x$ gives $B(x,x)B(z,x) = B(x,z)B(x,x)$ hence

$$B(x,x)\,(B(x,z) - B(z,x)) = 0 \quad (\forall x, z \in V).$$

We record these two observations in the form

$$\text{(II.A.7)} \qquad \begin{cases} \text{(i) } B(x,y)B(z,x) = B(x,z)B(y,x) \\ \text{(ii) } B(x,x) \neq 0 \implies B(x,z) = B(z,x) \ (\forall z). \end{cases}$$

If we are working over $\mathbb{R}$ or $\mathbb{C}$, at this point we are done: if some $B(x_0, x_0) \neq 0$ (i.e. B isn't alternating), then for any y, $B(x_0 + \epsilon y, x_0 + \epsilon y) \neq 0$ for small enough ϵ (by continuity). Applying (II.A.7)(ii) gives $B(x_0 + \epsilon y, z) = B(z, x_0 + \epsilon y)$ $(\forall z)$, whereupon differentiating $\frac{d}{d\epsilon}$ and setting $\epsilon = 0$ gives $B(y,z) = B(z,y)$ ($\implies$ B is symmetric).

For more general fields, we have to work harder. Suppose that B is neither symmetric ($B(u,v) \neq B(v,u)$ for some $u, v \in V$) nor alternating ($\exists w \in V$ with $B(w,w) \neq 0$). We will obtain a contradiction. Applying (II.A.7)(ii) yields

$$B(u,u) = 0 = B(v,v) \text{ and } \begin{cases} B(w,u) = B(u,w) \\ B(w,v) = B(v,w) \end{cases},$$

which together with (II.A.7)(i) gives

$$\begin{cases} B(u,v)B(w,u) = B(u,w)B(v,u) = B(w,u)B(v,u) \\ B(v,u)B(w,v) = B(v,w)B(u,v) = B(w,v)B(u,v) \end{cases}$$

and thus (using $B(u,v) \neq B(v,u)$ again) $B(w,u) = 0 = B(w,v)$. Since this also gives $B(u,w) = 0 = B(v,w)$, we can write

$$B(u, v+w) = B(u,v) + B(u,w) = B(u,v) \neq$$
$$B(v, yu) = B(v,u) + B(w,u) = B(v+w, u).$$

Applying (II.A.7)(ii) one more time, we see that $B(v+w, v+w) = 0$. But expanding this reveals $B(w,w) = 0$, a contradiction. □

II.A.8. DEFINITION. Let B and B' be bilinear forms on $\mathbb{F}$-vector spaces V and V'. We will call them *equivalent*, and write $B \sim B'$, if there exists an isomorphism $\mu\colon V \to V'$ such that the composition

$$V \otimes V \xrightarrow{\mu\otimes\mu} V' \otimes V' \xrightarrow{B'} \mathbb{F}$$

recovers B. (Clearly this means that if we write e' for the basis $\mu(e)$ of V', then $[B']_{e'} = [B]_e$.)

II.A.9. PROPOSITION-DEFINITION. *Suppose $B, \tilde{B}$ are two bilinear forms on V, and e any basis of V. Then $B \sim \tilde{B}$ if and only if $[B]_e$ and $[\tilde{B}]_e$ are **cogredient**, i.e. $[B]_e = {}^tS[\tilde{B}]_e S$ for some invertible matrix S.*

PROOF. If $B \sim \tilde{B}$, then there is an isomorphism $\mu\colon V \to V$ such that $[B]_e = [\tilde{B}]_{\tilde{e}}$, where $\tilde{e} = \mu(e)$. But $[\tilde{B}]_{\tilde{e}} = {}^tS[\tilde{B}]_e S$, where $S = {}_e[\mathbf{1}]_{\tilde{e}} = [\mu]_e$.

Conversely, since any invertible matrix S is $[\mu]_e$ (hence ${}_e[\mathbf{1}]_{\mu(e)}$) for some isomorphism μ, we get $[B]_e = [\tilde{B}]_{\mu(e)}$ hence $B \sim \tilde{B}$ (more precisely $B(x,y) = \tilde{B}(\mu(x), \mu(y))$). □