Class schedule: Bridge Course (Linear Algebra), Summer 2026

All slides, papers, and book chapters will be made available on the course webpage.

WEEK 1

Wed. July 1. Welcome, tea/coffee & donuts, introductions, course overview, discussion about our graduate program, future plans, etc.

Thurs. July 2. *Systems of equations and matrix multiplication.* Four ways to think about the fundamental problem in linear algebra: solving linear equations in n variables.

(1) Matrix form
(2) Row picture
(3) Column picture
(4) Grid picture

The example of $\{2x - y = 0, -x + 2y = 3\}$ was used. Then we did a 3×3 example: $\{2x - y = 0, -x + 2y - z = -1, -3y + 4z = 4\}$. Note how changing up the RHS changes the planes (row picture), but barely changes the column or grid pictures.

Then, we talked about 4 ways to multiply matrices $AB = C$, where A is $m \times n$ and B is $n \times p$.

(1) *Rows times columns*: $C_{ij} = (\text{row } i) \cdot (\text{column } j) = \sum_{k=1}^{n} a_{ik}b_{kj}$.
(2) *By columns*: $A[b_1 \cdots b_p] = [Ab_1 \cdots Ab_p]$. Each column Ab_i is a linear combination of the columns of A.
(3) *By rows*: $\begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} B = \begin{bmatrix} a_1^T B \\ \vdots \\ a_m^T B \end{bmatrix}$. Each row $a_j^T B$ is a linear combination of the rows of B.
(4) *Columns times rows.* This is a sum of n^2 rank-1 matrices, $\sum_{j,k=1}^{n} a_j b_k^T$.

Fri. July 3. The *four fundamental subspaces.* Given an $m \times n$ matrix A, we introduced the subspaces

(1) *Column space* $C(A)$ in $\mathbb{R}^m$
(2) *Row space* $C(A^T)$ in $\mathbb{R}^n$
(3) *Nullspace* $N(A)$ in $\mathbb{R}^n$
(4) *Left nullspace* $C(A^T)$ in $\mathbb{R}^m$.

We stated, without proving (will do later, in more generality) that $C(A)$ and $C(A^T)$ have the same rank, and that these subspaces come in orthogonal complement pairs:

$$\mathbb{R}^n = C(A) \oplus N(A^T), \qquad \mathbb{R}^m = C(A^T) \oplus N(A).$$

We proved that $N(A) = N(A^T A)$ under "undergraduate notation" (dot products). Then, we discussed inner products, and used $(x, y) = y^T Ax$ as an example, for $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. It was slightly cleaner to prove $N(A) = N(A^T A)$ with this notation.

We discussed subspaces and their sums, such as the difference between $Y + Z$ and $Y \oplus Z$. The former can *always* be defined, whereas the latter occurs only when $Y \cap Z = \{0\}$. An equivalent condition (HW) is that every element in $Y \oplus Z$ can be written uniquely as $y + z$, for $y \in Y$ and $z \in Z$. This lead to the formula

$$\dim(X + Y) = \dim Y + \dim Z + \dim(Y \cap Z),$$

with the caveat that (i) we haven't yet defined dimension, and (ii) we don't yet know how to prove this. We illustrated this with two examples: $Y = xy$-plane and $Z = yz$-plane, and also with

$Z = z$-axis.

We discussed how to solve an inhomogeneous system $Ax = b$. The "general solution" has the form $x = x_n + x_p$, where $Ax_n = 0$ and x_p is "any particular solution."

WEEK 2

Mon. July 6. *Projections and least squuares.* We started with an application of how the solution to the system $Ax = b$ has the form $x = x_n + x_p$: solving the ODE $y'' + 4y = 8$. This is the inhomomgeneous equation $Ly = 8$, for the linear operator $L = \frac{d^2}{dt^2}$. More generally, the nullspace of an n^{th} order linear differential operation is n-dimensional.

Then, we explored how to project a vector b onto another vector a. The result $p = xa$ satisfies $b = p + e$, and we derived $x = (a^T b)/(a^T a)$. Alternatively, we could describe $b \overset{P}{\longmapsto} p$ with the rank 1 matrix $P = (aa^T)/(a^T a)$. This matrix satisfies $P^T = P$ and $P^2 = P$, which are defining properties of projection matrices.

More generally, projections arise if we want to solve an underdetermined system $Ax = b$, where $b \notin C(A)$. The "best fit" solution is to solve $A\hat{x} = p$, where p is the projection of b onto $C(A)$. We showed how can be done by solving $A^T A\hat{x} = A^T b$. Reason: If $A = [a_1 \cdots a_n]$ and we write $A\hat{x} = p = \hat{x}_1 a_1 + \cdots + \hat{x}_n a_n$, then $a_i \perp e = (b - A\hat{x})$ means that $a_i^T (b - A\hat{x}) = 0$, which gives the equation $A^T (b - A\hat{x}) = 0$.

As a takeaway message, if S is a subspace with basis $a_1, \dots, a_r$, then the projection matrix onto SS is $P = A(A^T A)^{-1} A^T$, where $A = [a_1 \cdots a_r]$.

We finished by showing how a classic least squares problems can be viewed as an overdetermined linear system, using an example with three points $(1, 1)$, $(2, 2)$, and $(3, 2)$. Any line $b = C + Dt$ through these points would lead to a linear system $\{C + D = 1,\ C + 2D = 2,\ C + 3D = 2\}$ that has no solution. However, by solving $A^T A\hat{x} = A^T b$ instead, we find that the best fit line is when $C = 2/3$ and $D = 1/2$. Remark: we *can* find a degree-2 polynomial $b = C + Dt + Et^2$ that fits this data, because this would lead to a 3×3 system, which has a unique solution.

Tues. July 7 (2 classes). *Orthogonality, least squares, and QR factorization.* We reviewed what it means for a set of vectors to be orthogonal, and orthonormal. A square matrix Q is orthogonal if $Q^T Q = I$, which is equivalent to its columns being orthonormal.

Next, we discussed how to decompose a vector into an orthogonal basis. As an example, note that

$$v = (4, 3) = 4e_1 + 3e_2 = (v \cdot e_1)e_1 + (v \cdot e_2)e_2.$$

If we use a different basis, like $v_1 = (\sqrt{2}/2, \sqrt{2}/2)$ and $v_2 = (\sqrt{2}/2, \sqrt{2}/2)$.

$$v = (v \cdot v_1)v_1 + (v \cdot v_2)v_2 = 4.95v_1 + 0.701v_2.$$

We can do this with an orthogonal basis $w_1, \dots, w_n$ that isn't orthonormal by replacing $v \cdot w_i$ with $(v \cdot w_i)/(w_i \cdot w_i)$.

An example application of this is Fourier series. If $f(x)$ is a piecewise continuous 2π-periodic function, then it can be decomposed uniquely into a sum $f(x) = \frac{a_0}{2} + \sum a_n \cos(nx) + b_n \sin(nx)$. There are some technical details that require analysis to formalized, such as what happens at the points of discontinunity, an the fact that infinite sums are allowed. This all works because the set $\{\frac{1}{\sqrt{2}}, \cos(nx), \sin(nx) \mid n \in \mathbb{Z}\}$ is an orthonormal basis with respect to the inner product $\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\, dx$. We discussed complex Fourier series, and the original motivation for their discovery: solving PDEs on bounded domains.

Vector spaces and subspaces. We started with the Gram-Schmidt process, which takes an independent set of vectors, and outputs an orthonormal set. Finally, we mentioned how the Gram-Schmidt process can be described in matrix language. Specificaly, if $M = [a_1, \dots, a_n]$ is the original basis, and $Q = [q_1, \dots, q_n]$ the orthonormal basis from computing Gram-Schmidt, then these are related by $M = QR$, where $r_{ij} = q_i \cdot a_j$.

We gave the formal definition of a vector space. For this, we needed the formal definition of a group and a field. We discussed how to formally prove a few "obvious" facts, such as uniqueness of an identity element, uniqueness of inverses, and that $0x = \mathbf{0}$ for all vectors $x \in X$. We formally defined linear maps in this setting, though we discusse linearity earlier.

Wed. July 8. We defined linear combinations and spanning sets, and discussed the proof of

$$\mathrm{Span}(S) = \bigcap_{S \subseteq Y_\alpha \leq X} Y_\alpha.$$

The RHS of this is the "smallest subspace that contains S." It shows how a subspace can be defined "from the bottom, up" (as linear combinations), or "from the top, down." (as intersections)

We started with some examples of mathematical objects that can be defined "from the bottom, up" or "from the top, down." The convex hull is one such example. This equality fails for the defintion of an ideal, if the ring does not have unity (example, $I = (2)$ in $R = 2\mathbb{Z}$. In some cases, like the "smallest normal subgroup containing a set," there isn't a bottom-up definition.

Next, we talked about spanning sets, linear independence, and bases. We proved some basic facts, like how all bases of the same size, called its *dimension*, and gave examples of how this fails in groups (e.g., S_n). We showed how one can always extend an independent set to a basis, and how one can always remove vectors from a dependent set until it is independent.

Thurs. July 9. No class (make-up was Fri. July 3)

Fri. July 10. No class (make-up was Tues. July 7)

WEEK 3

Mon. July 13. We talked about complements and direct sums, and compared direct sums to direct products. These are the same when things are finite, but they differ in the infinite-dimensional cases. As an example, we compared the space $R \times R \times \cdots$ of infinite sequences ($\cong$ power series) to the space $R \otimes R \otimes \cdots$ of finite sums ($\cong$ polynomials). Just for fun, we discussed ℓ^p spaces from functional analysis. We also gave a sneak preview of the tensor product of two spaces, and how that differs from the direct product—dimension is multiplicative rather than additive. Finally, we defined what it meant for two vectors to be equivalent modulo Y.

Tues. July 14. Our two running examples for equivalence and quotient spaces used $Y = xy$-plane and $Z = z$-axis. We discussed what it meant to be well-defined, and left the proof that addition and scalar multiplication in X/Y being well-defined for the HW.

We proved that $\dim X = \dim Y + \dim X/Y$ and outlined how that can be used to prove

$$\dim U + \dim V - \dim(U \cap V) = \dim X.$$

The first step was to consider the case when $U \cap V = \{0\}$, i.e., $X = U \oplus V$, which we've already done. More generally, if $W = U \cap V$, when we can take the quotient of everything by W, and get that $\overline{X} = \overline{U} + \overline{V}$. This, with the previous results proven today, gives the desired result.

Wed. July 15. We spent this day going over some of the main ideas in group theory, with a focus on visualizations. Specifically, we discussed cosets, and how a quotient group G/H is defined iff and only if H is normal. Specifically, what it mean to be well-defined, and how and why can this fail? We focused on the isomorphism theorems, because they have analogues in linear algebra.

The first isomorphism theorem says that "*every homomorphic image is a quotient.*" Once we know that, the next natural question to ask is "*what are the substructures of the quotients?*," which is the second isomorphism theorem (correspondnence theorem). Then we can ask "*what is the result of taking the quotient by one of these substructures?*," which is the third isomorphism theorem (fraction theorem).

Thurs. July 16. We discussed how every group is a quotient of a free group, and how the construction is done by starting with a free group on the same number of generators, and taking the quotient of the smallest normal subgroup that contain all relators. Then, we introduced the idea of a univeral property, and how many familiar concepts can be formalized in this way, especially when speaking of "the largest" or "the smallest" object with a particular property. We stated and proved the universal property of quotients.

Fri. July 17. We abstracted the main ideas from the previous day to a "universal pair," and univeral mapping theorem. We showed how products and coproducts can be formalized via a universal mapping property. Then, we introduced the notion of a category, gave several examples, and finished with an overview of the fundamental group of a topological space, and the functor from the category of topological spaces to the category of groups.

Mon. July 20. We moved back to linear algebra and linear maps. We applied the first isomorphism theorem ($X/\operatorname{Ker} f \cong \operatorname{Im} f$) to $\dim Y + \dim X/Y = \dim X$ to get the rank-nullity theorem: $\dim X = \operatorname{rank} f + \operatorname{nullity} f$. We took a tangent and showed two applications of the rank-nullity theorem to polynomialsi: interpolation and average values over intervals.

Then, we discussed the dual space X', which is the space of all linear scalar functions $\ell \colon X \to K$. We showed how every co-vector can be written as $\ell(x) = a_1c_1 + \cdots + a_nc_n$, where $x = a_1x_1 + \cdots + a_nx_n$. In light of this, one can (and should!) think of elements in X (vectors) as column vectors, and elements in X' (co-vectors) as row vectors. It is convenient to denote $(\ell, x) := \ell(x)$. Given a basis $x_1, \dots, x_n$, the *dual* basis in X' is $\ell_1, \dots, \ell_n$, where $\ell_i(x_j) = \delta_{ij}$.

We reviewed the dual space and the advantage of using the notation $(\ell, x) = \ell(x)$. This helped us understand the double dual X'', whose elements can be thought of as "evaluation maps." The space X'' can be canonically identified with X, whereas there is no such basis-free identification with X'.

Tues. July 21. We defined the *annihilator* $Y^\perp$ of a subspace $Y \leq X$, and proved a few basic properties about it. We did this so we could define the *transpose* of a linear map $T \colon X \to U$. This is a map $T \colon U' \to X'$ such that $(\ell, Tx) = (T'\ell, x)$ for all $x \in X$ and $u \in U$. We gave short proofs of some basic properties, such as $(TS)' = S'T'$, $R^\perp = N_{T'}$, and $\operatorname{rank}(T) = \operatorname{rank}(T')$. All of these have special cases in terms of matrices that are *much* harder to prove directly.

The *Riesz representation theorem* from functional analysis "basically" says that in a Hilbert space, every linear functional is the the inner product on some fixed vector (there are other technical conditions). For a non-example, note that in $\ell_1(\mathbb{R})$, the "dot product with $(1, 1, 1, \dots)$" is a linear functional, despite $(1, 1, 1, \dots) \notin \ell_1(\mathbb{R})$. We discussed the difference between the transpose of a linear (a basis-free, and inner product free concept), and the adjoint, which depends on an inner product and used the Riesz representation theorem.