

8.5 Fourier Series: Linear Algebra for Functions

This section goes from finite dimensions to *infinite* dimensions. I want to explain linear algebra in infinite-dimensional space, and to show that it still works. First step: look back. This book began with vectors and dot products and linear combinations. We begin by converting those basic ideas to the infinite case—then the rest will follow.

What does it mean for a vector to have infinitely many components? There are two different answers, both good:

1. The vector becomes $\mathbf{v} = (v_1, v_2, v_3, \dots)$. It could be $(1, \frac{1}{2}, \frac{1}{4}, \dots)$.
2. The vector becomes a function $f(x)$. It could be $\sin x$.

We will go both ways. Then the idea of Fourier series will connect them.

After vectors come *dot products*. The natural dot product of two infinite vectors (v_1, v_2, \dots) and (w_1, w_2, \dots) is an infinite series:

$$\text{Dot product} \quad \mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots \quad (1)$$

This brings a new question, which never occurred to us for vectors in \mathbf{R}^n . Does this infinite sum add up to a finite number? Does the series converge? Here is the first and biggest difference between finite and infinite.

When $\mathbf{v} = \mathbf{w} = (1, 1, 1, \dots)$, the sum certainly does not converge. In that case $\mathbf{v} \cdot \mathbf{w} = 1 + 1 + 1 + \dots$ is infinite. Since \mathbf{v} equals \mathbf{w} , we are really computing $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2 = \text{length squared}$. The vector $(1, 1, 1, \dots)$ has infinite length. *We don't want that vector.* Since we are making the rules, we don't have to include it. The only vectors to be allowed are those with finite length:

DEFINITION The vector (v_1, v_2, \dots) is in our infinite-dimensional “*Hilbert space*” if and only if its length $\|\mathbf{v}\|$ is finite:

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + v_3^2 + \dots \text{ must add to a finite number.}$$

Example 1 The vector $\mathbf{v} = (1, \frac{1}{2}, \frac{1}{4}, \dots)$ is included in Hilbert space, because its length is $2/\sqrt{3}$. We have a geometric series that adds to $4/3$. The length of \mathbf{v} is the square root:

$$\text{Length squared} \quad \mathbf{v} \cdot \mathbf{v} = 1 + \frac{1}{4} + \frac{1}{16} + \dots = \frac{1}{1 - \frac{1}{4}} = \frac{4}{3}.$$

If \mathbf{v} and \mathbf{w} have finite length, how large can their dot product be?

Question If \mathbf{v} and \mathbf{w} have finite length, how large can their dot product be?

Answer The sum $\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots$ also adds to a finite number. We can safely take dot products. The Schwarz inequality is still true:

$$\text{Schwarz inequality} \quad |\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|. \quad (2)$$

The ratio of $\mathbf{v} \cdot \mathbf{w}$ to $\|\mathbf{v}\| \|\mathbf{w}\|$ is still the cosine of θ (the angle between \mathbf{v} and \mathbf{w}). Even in infinite-dimensional space, $|\cos \theta|$ is not greater than 1.

Now change over to functions. Those are the “vectors.” The space of functions $f(x)$, $g(x)$, $h(x)$, . . . defined for $0 \leq x \leq 2\pi$ must be somehow bigger than \mathbf{R}^n . **What is the dot product of $f(x)$ and $g(x)$? What is the length of $f(x)$?**

Key point in the continuous case: *Sums are replaced by integrals.* Instead of a sum of v_j times w_j , the dot product is an integral of $f(x)$ times $g(x)$. Change the “dot” to parentheses with a comma, and change the words “dot product” to *inner product*:

DEFINITION The *inner product* of $f(x)$ and $g(x)$, and the *length squared*, are

$$(f, g) = \int_0^{2\pi} f(x)g(x) dx \quad \text{and} \quad \|f\|^2 = \int_0^{2\pi} (f(x))^2 dx. \quad (3)$$

The interval $[0, 2\pi]$ where the functions are defined could change to a different interval like $[0, 1]$ or $(-\infty, \infty)$. We chose 2π because our first examples are $\sin x$ and $\cos x$.

Example 2 The length of $f(x) = \sin x$ comes from its inner product with itself:

$$(f, f) = \int_0^{2\pi} (\sin x)^2 dx = \pi. \quad \text{The length of } \sin x \text{ is } \sqrt{\pi}.$$

That is a standard integral in calculus—not part of linear algebra. By writing $\sin^2 x$ as $\frac{1}{2} - \frac{1}{2} \cos 2x$, we see it go above and below its average value $\frac{1}{2}$. Multiply that average by the interval length 2π to get the answer π .

More important: $\sin x$ and $\cos x$ are *orthogonal in function space*:

$$\begin{array}{ll} \text{Inner product} & \int_0^{2\pi} \sin x \cos x dx = \int_0^{2\pi} \frac{1}{2} \sin 2x dx = \left[-\frac{1}{4} \cos 2x \right]_0^{2\pi} = 0. \end{array} \quad (4)$$

This zero is no accident. It is highly important to science. The orthogonality goes beyond the two functions $\sin x$ and $\cos x$, to an infinite list of sines and cosines. The list contains $\cos 0x$ (which is 1), $\sin x$, $\cos x$, $\sin 2x$, $\cos 2x$, $\sin 3x$, $\cos 3x$, . . .

Every function in that list is orthogonal to every other function in the list.

Fourier Series

The Fourier series of a function $y(x)$ is its expansion into sines and cosines:

$$y(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots \quad (5)$$

We have an orthogonal basis! The vectors in “function space” are combinations of the sines and cosines. On the interval from $x = 2\pi$ to $x = 4\pi$, all our functions repeat what they did from 0 to 2π . They are “*periodic*.” The distance between repetitions is the period 2π .

Remember: The list is infinite. The Fourier series is an infinite series. We avoided the vector $v = (1, 1, 1, \dots)$ because its length is infinite, now we avoid a function like $\frac{1}{2} + \cos x + \cos 2x + \cos 3x + \dots$. (Note: This is π times the famous **delta function** $\delta(x)$. It is an infinite “spike” above a single point. At $x = 0$ its height $\frac{1}{2} + 1 + 1 + \dots$ is infinite. At all points inside $0 < x < 2\pi$ the series adds in some average way to zero.) The integral of $\delta(x)$ is 1. But $\int \delta^2(x) = \infty$, so delta functions are excluded from Hilbert space.

Compute the length of a typical sum $f(x)$:

$$\begin{aligned}(f, f) &= \int_0^{2\pi} (a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + \dots)^2 dx \\ &= \int_0^{2\pi} (a_0^2 + a_1^2 \cos^2 x + b_1^2 \sin^2 x + a_2^2 \cos^2 2x + \dots) dx \\ \|f\|^2 &= 2\pi a_0^2 + \pi(a_1^2 + b_1^2 + a_2^2 + \dots).\end{aligned}\tag{6}$$

The step from line 1 to line 2 used orthogonality. All products like $\cos x \cos 2x$ integrate to give zero. Line 2 contains what is left—the integrals of each sine and cosine squared. Line 3 evaluates those integrals. (The integral of 1^2 is 2π , when all other integrals give π .) If we divide by their lengths, our functions become *orthonormal*:

$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots$ is an orthonormal basis for our function space.

These are unit vectors. We could combine them with coefficients $A_0, A_1, B_1, A_2, \dots$ to yield a function $F(x)$. Then the 2π and the π 's drop out of the formula for length.

$$\text{Function length} = \text{vector length} \quad \|F\|^2 = (F, F) = A_0^2 + A_1^2 + B_1^2 + A_2^2 + \dots \tag{7}$$

Here is the important point, for $f(x)$ as well as $F(x)$. *The function has finite length exactly when the vector of coefficients has finite length.* Fourier series gives us a perfect match between function space and infinite-dimensional Hilbert space. The function is in L^2 , its Fourier coefficients are in ℓ^2 .

The function space contains $f(x)$ exactly when the Hilbert space contains the vector $v = (a_0, a_1, b_1, \dots)$ of Fourier coefficients. Both $f(x)$ and v have finite length.

Example 3 Suppose $f(x)$ is a “square wave,” equal to 1 for $0 \leq x < \pi$. Then $f(x)$ drops to -1 for $\pi \leq x < 2\pi$. The $+1$ and -1 repeats forever. This $f(x)$ is an odd function like the sines, and all its cosine coefficients are zero. We will find its Fourier series, containing only sines:

$$\text{Square wave} \quad f(x) = \frac{4}{\pi} \left[\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \dots \right] \tag{8}$$

The length is $\sqrt{2\pi}$, because at every point $(f(x))^2$ is $(-1)^2$ or $(+1)^2$:

$$\|f\|^2 = \int_0^{2\pi} (f(x))^2 dx = \int_0^{2\pi} 1 dx = 2\pi.$$

At $x = 0$ the sines are zero and the Fourier series gives zero. This is half way up the jump from -1 to $+1$. The Fourier series is also interesting when $x = \frac{\pi}{2}$. At this point the square wave equals 1 , and the sines in (8) alternate between $+1$ and -1 :

$$\text{Formula for } \pi \quad 1 = \frac{4}{\pi} \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \right). \quad (9)$$

Multiply by π to find a magical formula $4(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots)$ for that famous number.

The Fourier Coefficients

How do we find the a 's and b 's which multiply the cosines and sines? For a given function $f(x)$, we are asking for its Fourier coefficients:

$$\text{Fourier series} \quad f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + \dots.$$

Here is the way to find a_1 . **Multiply both sides by $\cos x$. Then integrate from 0 to 2π .** The key is orthogonality! All integrals on the right side are zero, except for $\cos^2 x$:

$$\text{Coefficient } a_1 \quad \int_0^{2\pi} f(x) \cos x \, dx = \int_0^{2\pi} a_1 \cos^2 x \, dx = \pi a_1. \quad (10)$$

Divide by π and you have a_1 . To find any other a_k , multiply the Fourier series by $\cos kx$. Integrate from 0 to 2π . Use orthogonality, so only the integral of $a_k \cos^2 kx$ is left. That integral is πa_k , and divide by π :

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx \quad \text{and similarly} \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx. \quad (11)$$

The exception is a_0 . This time we multiply by $\cos 0x = 1$. The integral of 1 is 2π :

$$\text{Constant term} \quad a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) \cdot 1 \, dx = \text{average value of } f(x). \quad (12)$$

I used those formulas to find the Fourier coefficients for the square wave. The integral of $f(x) \cos kx$ was zero. The integral of $f(x) \sin kx$ was $4/k$ for odd k .

Compare Linear Algebra in \mathbb{R}^n

The point to emphasize is how this infinite-dimensional case is so much like the n -dimensional case. Suppose the nonzero vectors v_1, \dots, v_n are orthogonal. We want to write the vector b (instead of the function $f(x)$) as a combination of those v 's:

$$\text{Finite orthogonal series} \quad b = c_1 v_1 + c_2 v_2 + \dots + c_n v_n. \quad (13)$$

Multiply both sides by v_1^T . Use orthogonality, so $v_1^T v_2 = 0$. Only the c_1 term is left:

$$\text{Coefficient } c_1 \quad v_1^T b = c_1 v_1^T v_1 + 0 + \dots + 0. \quad \text{Therefore } c_1 = v_1^T b / v_1^T v_1. \quad (14)$$

The denominator $v_1^T v_1$ is the length squared, like π in equation 11. The numerator $v_1^T b$ is the inner product like $\int f(x) \cos kx \, dx$. **Coefficients are easy to find when the basis**

vectors are orthogonal. We are just doing one-dimensional projections, to find the components along each basis vector.

The formulas are even better when the vectors are orthonormal. Then we have unit vectors. The denominators $v_k^T v_k$ are all 1. You know $c_k = v_k^T b$ in another form:

$$\text{Equation for } c\text{'s} \quad c_1 v_1 + \cdots + c_n v_n = b \quad \text{or} \quad \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = b.$$

The v 's are in an orthogonal matrix Q . Its inverse is Q^T . That gives the c 's:

$$Qc = b \quad \text{yields} \quad c = Q^T b. \quad \text{Row by row this is } c_k = q_k^T b.$$

Fourier series is like having a matrix with infinitely many orthogonal columns. Those columns are the basis functions $1, \cos x, \sin x, \dots$. After dividing by their lengths we have an “infinite orthogonal matrix.” Its inverse is its transpose. Orthogonality is what reduces a series of terms to one single term.

Problem Set 8.5

- 1 Integrate the trig identity $2 \cos jx \cos kx = \cos(j+k)x + \cos(j-k)x$ to show that $\cos jx$ is orthogonal to $\cos kx$, provided $j \neq k$. What is the result when $j = k$?
- 2 Show that $1, x$, and $x^2 - \frac{1}{3}$ are orthogonal, when the integration is from $x = -1$ to $x = 1$. Write $f(x) = 2x^2$ as a combination of those orthogonal functions.
- 3 Find a vector (w_1, w_2, w_3, \dots) that is orthogonal to $v = (1, \frac{1}{2}, \frac{1}{4}, \dots)$. Compute its length $\|w\|$.
- 4 The first three *Legendre polynomials* are $1, x$, and $x^2 - \frac{1}{3}$. Choose c so that the fourth polynomial $x^3 - cx$ is orthogonal to the first three. All integrals go from -1 to 1 .
- 5 For the square wave $f(x)$ in Example 3, show that

$$\int_0^{2\pi} f(x) \cos x \, dx = 0 \quad \int_0^{2\pi} f(x) \sin x \, dx = 4 \quad \int_0^{2\pi} f(x) \sin 2x \, dx = 0.$$

Which three Fourier coefficients come from those integrals?

- 6 The square wave has $\|f\|^2 = 2\pi$. Then (6) gives what remarkable sum for π^2 ?
- 7 Graph the square wave. Then graph by hand the sum of two sine terms in its series, or graph by machine the sum of 2, 3, and 10 terms. The famous *Gibbs phenomenon* is the oscillation that overshoots the jump (this doesn't die down with more terms).
- 8 Find the lengths of these vectors in Hilbert space:

$$(a) \quad v = \left(\frac{1}{\sqrt{1}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{4}}, \dots \right)$$

8.6 Linear Algebra for Statistics and Probability

Statistics deals with data, often in large quantities. Since data tends to go into rectangular matrices, we expect to see $A^T A$. The least squares problem $A\hat{x} \approx b$ is **linear regression**. The best solution \hat{x} fits m observations by $n < m$ parameters. This is a fundamental application of linear algebra to statistics.

This section goes beyond $A^T A\hat{x} = A^T b$. These unweighted equations assume that the measurements b_1, \dots, b_m are equally reliable. When there is good reason to expect higher accuracy (lower variance) in some b_i , those equations should be weighted more heavily. With what weights w_1, \dots, w_m ? And if the b_i are not independent, a **covariance matrix** Σ gives the statistics of the errors. Here are key topics in this section:

1. Weighted least squares and $A^T C A \hat{x} = A^T C b$
2. Variances $\sigma_1^2, \dots, \sigma_m^2$ and the covariance matrix Σ
3. Important probability distributions: binomial, Poisson, and normal
4. Principal Component Analysis (PCA) to find combinations with greatest variance.

Weighted Least Squares

To include weights in the m equations $Ax = b$, multiply each equation i by a weight w_i . Put those m weights into a diagonal matrix W . We are replacing $Ax = b$ by $WAx = Wb$.

The equations are no more and no less solvable—we expect to use least squares.

The least squares equation $A^T A\hat{x} = A^T b$ changes to $(WA)^T WA\hat{x} = (WA)^T Wb$.

The matrix $C = W^T W$ is inside $(WA)^T WA$, in the middle of weighted least squares.

Weighted least squares $C = W^T W$ is in the n equations for \hat{x} $A^T C A \hat{x} = A^T C b$ (1)

When $n = 1$ and A = column of 1's, \hat{x} changes from an average to a weighted average:

Simplest case $\hat{x} = \frac{b_1 + \dots + b_m}{m}$ changes to $\hat{x}_W = \frac{w_1^2 b_1 + \dots + w_m^2 b_m}{w_1^2 + \dots + w_m^2}$. (2)

This average \hat{x}_W gives greatest weight to the observations b_i that have the largest w_i . We always assume that errors have zero mean. (Subtract the mean if necessary, so there is no one-sided bias in the measurements.)

How should we choose the weights w_i ? This depends on the reliability of b_i . If that observation has variance σ_i^2 , then the root mean square error in b_i is σ_i . When we divide the equations by $\sigma_1, \dots, \sigma_m$ (left side together with right side), all variances will equal 1. So the weight is $w_i = 1/\sigma_i$ and the diagonal of $C = W^T W$ contains the numbers $1/\sigma_i^2$.

The statistically correct matrix is $C = \text{diag}(1/\sigma_1^2, \dots, 1/\sigma_m^2)$.

This is correct provided the errors e_i and e_j in different equations are statistically independent. If the errors are dependent, off-diagonal entries show up in the covariance matrix Σ . The good choice is still $C = \Sigma^{-1}$ as described in this section.

Mean and Variance

The two crucial numbers for a random variable are its **mean** m and its **variance** σ^2 . The “expected value” $E[e]$ is found from the probabilities p_1, p_2, \dots of the possible errors e_1, e_2, \dots (and the variance σ^2 is always measured around the mean).

For a discrete random variable, the error e_j has probability p_j (the p_j add to 1):

$$\text{Mean } m = E[e] = \sum e_j p_j \quad \text{Variance } \sigma^2 = E[(e - m)^2] = \sum (e_j - m)^2 p_j \quad (3)$$

Example 1 Flip a fair coin. The result is 1 (for heads) or 0 (for tails). Those events have equal probabilities $p_0 = p_1 = 1/2$. The mean is $m = 1/2$ and the variance is $\sigma^2 = 1/4$:

$$\text{Mean} = (0) \frac{1}{2} + (1) \frac{1}{2} \quad \text{Variance} = \left(0 - \frac{1}{2}\right)^2 \frac{1}{2} + \left(1 - \frac{1}{2}\right)^2 \frac{1}{2} = \frac{1}{4}.$$

Example 2 (Binomial) Flip the fair coin N times and count heads. With 3 flips, we see $M = 0, 1, 2$, or 3 heads. The chances are $1/8, 3/8, 3/8, 1/8$. There are three ways to see $M = 2$ heads: HHT, HTH, and THH, and only HHH for $M = 3$ heads.

For all N , the number of ways to see M heads is the binomial coefficient “ N choose M ”. Divide by the total number 2^N of all possible outcomes to get the probability for each M :

$$\begin{matrix} M \text{ heads in} \\ N \text{ coin flips} \end{matrix} \quad p_M = \frac{1}{2^N} \binom{N}{M} = \frac{1}{2^N} \frac{N!}{M!(N-M)!} \quad \text{Check } \frac{1}{2^3} \frac{3!}{2!1!} = \frac{3}{8} \quad (4)$$

Gamblers know this instinctively. The probabilities p_M add to $(\frac{1}{2} + \frac{1}{2})^N = 1$. The mean value of the number of heads is $m = N/2$. The variance around m turns out to be $\sigma^2 = N/4$. The standard deviation $\sigma = \sqrt{N}/2$ measures the expected spread around the mean.

Example 3 (Poisson) A very unfair coin (small $p << \frac{1}{2}$) is flipped very often (large N). **The product** $\lambda = pN$ **is kept fixed**. The high probability of tails is $1 - p$ each time. So the chance p_0 of no heads in N flips (tails every time) is $(1 - p)^N = (1 - \lambda/N)^N$. For large N this approaches $e^{-\lambda}$. The probability p_j of j heads in N very unfair flips comes out neatly in terms of the crucial number $\lambda = pN$:

$$\text{Poisson probabilities} \quad p_j = \frac{\lambda^j}{j!} e^{-\lambda} \quad \text{Mean } m = \lambda \quad \text{Variance } \sigma^2 = \lambda \quad (5)$$

Poisson applies to counting infrequent events (low p) over a long time T . Then $\lambda = pT$.

A *continuous* random variable will have a probability *density* function $p(x)$ instead of p_1, p_2, \dots “An outcome between x and $x + dx$ has probability $p(x) dx$.” The total probability is $\int p(x) dx = 1$, since some outcome must happen. Sums become integrals:

$$\text{Mean } m = \text{Expected value} = \int x p(x) dx \quad \text{Variance } \sigma^2 = \int (x - m)^2 p(x) dx. \quad (6)$$

The outstanding example of a probability density function $p(x)$ (called the pdf) is the **normal distribution** $N(0, \sigma)$. This has mean zero by symmetry. Its variance is σ^2 :

Normal (Gaussian)
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \quad \text{with} \quad \int_{-\infty}^{\infty} p(x) dx = 1. \quad (7)$$

The graph of $p(x)$ is the famous bell-shaped curve. The integral of $p(x)$ from $-\sigma$ to σ is the probability that a random sample is less than one standard deviation σ from the mean. This is near 2/3. MATLAB's **randn** uses the normal distribution with $\sigma = 1$.

This normal $p(x)$ appears everywhere because of the **Central Limit Theorem**: The average over many independent trials of another distribution (like binomial) will approach a normal distribution as $N \rightarrow \infty$. A shift produces $m = 0$ and rescaling produces $\sigma = 1$.

Normalized headcount $x = \frac{M - \text{mean}}{\sigma} = \frac{M - N/2}{\sqrt{N}/2} \rightarrow \text{Normal } N(0, 1).$

The Covariance Matrix

Now run m different experiments at once. They might be independent, or there might be some correlation between them. Each measurement b is now a *vector* with m components. Those components are the outputs b_i from the m experiments.

If we measure distances from the means m_i , each error $e_i = b_i - m_i$ has *mean zero*. If two errors e_i and e_j are *independent* (no relation between them), their product $e_i e_j$ also has mean zero. But if the measurements are by the same observer at nearly the same time, the errors e_i and e_j could tend to have the same sign or the same size. **The errors in the m experiments could be correlated**. The products $e_i e_j$ are weighted by p_{ij} (their probability): covariance $\sigma_{ij} = \sum \sum p_{ij} e_i e_j$. The sum of $e_i^2 p_{ii}$ is the variance σ_i^2 :

Covariance $\sigma_{ij} = \sigma_{ji} = E[e_i e_j] = \text{expected value of } (e_i \text{ times } e_j). \quad (8)$

This is the (i, j) and (j, i) entry of the **covariance matrix** Σ . The (i, i) entry is $\sigma_{ii} = \sigma_i^2$.

Example 4 (Multivariate normal) For m random variables, the probability density function moves from $p(x)$ to $p(b) = p(b_1, \dots, b_m)$. The normal distribution with mean zero was controlled by one positive number σ^2 . Now $p(b)$ is controlled by an m by m positive definite matrix Σ . This is the covariance matrix and its determinant is $|\Sigma|$:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \quad \text{becomes} \quad p(b) = \frac{1}{(2\pi)^{m/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} b^T \Sigma^{-1} b} \quad (9)$$

The integral of $p(b)$ over m -dimensional space is 1. The integral of $bb^T p(b)$ is Σ .

The good way to handle that exponent $-\mathbf{b}^T \Sigma^{-1} \mathbf{b}/2$ is to use the eigenvalues and orthonormal eigenvectors of Σ (*linear algebra enters here*). When $\Sigma = Q\Lambda Q^T = Q\Lambda Q^{-1}$, replacing \mathbf{b} by $Q\mathbf{c}$ will split $p(\mathbf{b})$ into m one-dimensional normal distributions:

$$\exp(-\mathbf{b}^T \Sigma^{-1} \mathbf{b}/2) = \exp(-\mathbf{c}^T \Lambda^{-1} \mathbf{c}/2) = (e^{-c_1^2/2\lambda_1}) \cdots (e^{-c_m^2/2\lambda_m}).$$

The determinant has $|\Sigma|^{1/2} = |\Lambda|^{1/2} = (\lambda_1 \cdots \lambda_m)^{1/2}$. Each integral over $-\infty < c_i < \infty$ is back to one dimension, where $\lambda = \sigma^2$. Notice the wonderful fact that after any linear transformation (here $\mathbf{c} = Q^{-1}\mathbf{b}$), we still have a multivariate normal distribution.

We could even reach variances = 1 by including $\sqrt{\Lambda}$ in the change from \mathbf{b} to \mathbf{z} :

**Standard
normal** $\mathbf{b} = \sqrt{\Lambda} Q \mathbf{z}$ changes $p(\mathbf{b})d\mathbf{b}$ to $p(\mathbf{z})d\mathbf{z} = \frac{e^{-\mathbf{z}^T \mathbf{z}/2}}{(2\pi)^{m/2}} d\mathbf{z}$

This tells us the right weight matrix W to bring $A\mathbf{x} = \mathbf{b}$ back to ordinary least squares for $WA\mathbf{x} = W\mathbf{b}$. We want $W\mathbf{b}$ to become the standard normal \mathbf{z} . So W will be the inverse of $\sqrt{\Lambda} Q$. Better than that, $C = W^T W$ is the inverse of $Q\Lambda Q^T$ which is Σ .

Summary For independent errors, Σ is the diagonal matrix $\text{diag}(\sigma_1^2, \dots, \sigma_m^2)$. This is the usual choice. The right weights w_i for the equations $A\mathbf{x} = \mathbf{b}$ are $1/\sigma_1, \dots, 1/\sigma_m$ (this will equalize all variances to 1). The right matrix $C = W^T W$ in the middle of the weighted least squares equations is exactly Σ^{-1} :

Weighted least squares $A^T \Sigma^{-1} A \hat{\mathbf{x}} = A^T \Sigma^{-1} \mathbf{b}$ (9)

This choice of weighting returns $A\mathbf{x} = \mathbf{b}$ to a least squares problem $WA\mathbf{x} = W\mathbf{b}$ with equally reliable and independent errors. The usual equation $(WA)^T WA \hat{\mathbf{x}} = (WA)^T W \mathbf{b}$ is the same as (9).

It was Gauss who found this *best linear unbiased estimate* $\hat{\mathbf{x}}$. Unbiased because the mean of $\mathbf{x} - \hat{\mathbf{x}}$ is zero, linear because of equation (9), best because the covariance of $\mathbf{x} - \hat{\mathbf{x}}$ is as small as possible. That covariance (for error in $\hat{\mathbf{x}}$, not error in \mathbf{b} !) is important:

Covariance of the best $\hat{\mathbf{x}}$ $P = E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] = (A^T \Sigma^{-1} A)^{-1}$ (10)

Example 5 Your pulse rate is measured ten times by independent doctors, all equally reliable. The mean error of each b_i is zero, and each variance is σ^2 . Then $\Sigma = \sigma^2 I$. The ten equations $x = b_i$ produce the 10 by 1 matrix A of all ones. The best estimate $\hat{\mathbf{x}}$ is the average of the ten b_i . The variance of that average value $\hat{\mathbf{x}}$ is the number P :

$$P = (A^T \Sigma^{-1} A)^{-1} = \sigma^2 / 10 \quad \text{so averaging reduces the variance.}$$

This matrix $P = (A^T \Sigma^{-1} A)^{-1}$ tells how reliable is the result $\hat{\mathbf{x}}$ of the experiment (Problem 6). P does not depend on the b 's in the actual experiment! Those b 's have probability distributions. Each experiment produces a sample value of $\hat{\mathbf{x}}$ from a sample b .

When a small Σ gives good reliability of the inputs b , a small P gives good reliability of the outputs \hat{x} . The key formula $P = (A^T \Sigma^{-1} A)^{-1}$ connects those covariances.

Principal Component Analysis

These paragraphs are about finding useful information in a data matrix A . Start by measuring m properties (m features) of n samples. These could be grades in m courses for n students (a row for each course, a column for each student). From each row, subtract its average so the sample means are zero. We look for a **combination of courses** and/or **combination of students** for which the data provides the most information.

Information is “distance from randomness” and it is measured by **variance**. A large variance in course grades means greater information than a small variance.

The key matrix idea is the Singular Value Decomposition $A = U \Sigma V^T$. We are back again to $A^T A$ and $A A^T$, because their unit eigenvectors are the singular vectors v_1, \dots, v_n in V and u_1, \dots, u_m in U . The singular values in the diagonal matrix Σ (not the covariance) are in decreasing order and σ_1 is the most important. Weighting the m courses by the components of u_1 gives a “**master course**” or “**eigencourse**” with the most significant grades.

Example 6 Suppose the grades A, B, C, F are worth 4, 2, 0, -6 points. If each course and each student has one of each grade, then all means are zero. Here is the grade matrix A with (1, 1, 1, 1) in its nullspace (rank 3). To keep integers, the SVD of A will be written as $2U$ times $\Sigma/4$ times $(2V)^T$. So the σ ’s are 12, 8, 4:

$$\begin{bmatrix} -6 & 2 & 0 & 4 \\ 0 & 4 & -6 & 2 \\ 4 & 0 & 2 & -6 \\ 2 & -6 & 4 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & & & \\ & 2 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Weighting the rows (the courses) by $u_1 = \frac{1}{2}(-1, -1, 1, 1)$ will give the *eigencourse*. Weighting the columns (the students) by $v_1 = \frac{1}{2}(1, -1, 1, -1)$ gives the *eigenstudent*. Weighting the columns (the students) by $v_1 = \frac{1}{2}(1, -1, 1, -1)$ gives the *eigenstudent*. The fraction of the grade matrix that is “explained” by that one course and student is $\sigma_1^2 / (\sigma_1^2 + \sigma_2^2 + \sigma_3^2) = 9/14$. The σ ’s in the SVD are the variances σ^2 .

I guess this master course is what a Director of Admissions is looking for. If all grades in gym are the same, that row of A will be all zero—and gym is not part of the master course. Probably calculus is a part, but what about students who don’t take calculus? The problem of **missing data** (holes in the matrix A) is extremely difficult for social sciences and the census and so much of the statistics of experiments.

Gene expression data Determining the functions of genes, and combinations of genes, is a central problem of genetics. Which genes combine to give which properties? Which genes malfunction to give which diseases?

We now have an incredibly fast way to find gene expression data in the lab. A **microarray** is often packed onto an Affymetrix chip, measuring tens of thousands of genes from one sample (one person). The understanding of genetic data (*bioinformatics*) has become a tremendous application of linear algebra.

8.7 Computer Graphics

Computer graphics deals with images. The images are moved around. Their scale is changed. Three dimensions are projected onto two dimensions. All the main operations are done by matrices—but the shape of these matrices is surprising.

The transformations of three-dimensional space are done with 4 by 4 matrices. You would expect 3 by 3. The reason for the change is that one of the four key operations cannot be done with a 3 by 3 matrix multiplication. Here are the four operations:

Translation (shift the origin to another point $P_0 = (x_0, y_0, z_0)$)

Rescaling (by c in all directions or by different factors c_1, c_2, c_3)

Rotation (around an axis through the origin or an axis through P_0)

Projection (onto a plane through the origin or a plane through P_0).

Translation is the easiest—just add (x_0, y_0, z_0) to every point. But this is not linear! No 3 by 3 matrix can move the origin. So we change the coordinates of the origin to $(0, 0, 0, 1)$. This is why the matrices are 4 by 4. The “*homogeneous coordinates*” of the point (x, y, z) are $(x, y, z, 1)$ and we now show how they work.

1. Translation Shift the whole three-dimensional space along the vector v_0 . The origin moves to (x_0, y_0, z_0) . This vector v_0 is added to every point v in \mathbf{R}^3 . Using homogeneous coordinates, the 4 by 4 matrix T shifts the whole space by v_0 :

$$\text{Translation matrix } T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ x_0 & y_0 & z_0 & 1 \end{bmatrix}.$$

Important: *Computer graphics works with row vectors.* We have row times matrix instead of matrix times column. You can quickly check that $[0 \ 0 \ 0 \ 1] T = [x_0 \ y_0 \ z_0 \ 1]$.

To move the points $(0, 0, 0)$ and (x, y, z) by v_0 , change to homogeneous coordinates $(0, 0, 0, 1)$ and $(x, y, z, 1)$. Then multiply by T . A row vector times T gives a row vector.

$(0, 0, 0, 1)$ and $(x + x_0, y + y_0, z + z_0, 1)$.

Every v moves to $v + v_0$: $[x \ y \ z \ 1] T = [x + x_0 \ y + y_0 \ z + z_0 \ 1]$.

The output tells where any v will move. (It goes to $v + v_0$.) Translation is now achieved by a matrix, which was impossible in \mathbf{R}^3 .

2. Scaling To make a picture fit a page, we change its width and height. A Xerox copier will rescale a figure by 90%. In linear algebra, we multiply by .9 times the identity matrix. That matrix is normally 2 by 2 for a plane and 3 by 3 for a solid. In computer graphics, with homogeneous coordinates, the matrix is *one size larger*:

$$\text{Rescale the plane: } S = \begin{bmatrix} .9 & & & \\ & .9 & & \\ & & 1 & \end{bmatrix}$$

$$\text{Rescale a solid: } S = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Important: S is not cI . We keep the “1” in the lower corner. Then $[x, y, 1]$ times S is the correct answer in homogeneous coordinates. The origin stays in its normal position because $[0 \ 0 \ 1]S = [0 \ 0 \ 1]$.

If we change that 1 to c , the result is strange. **The point** (cx, cy, cz, c) is the same as $(x, y, z, 1)$. The special property of homogeneous coordinates is that multiplying by cI does not move the point. The origin in \mathbf{R}^3 has homogeneous coordinates $(0, 0, 0, 1)$ and $(0, 0, 0, c)$ for every nonzero c . This is the idea behind the word “homogeneous.”

Scaling can be different in different directions. To fit a full-page picture onto a half-page, scale the y direction by $\frac{1}{2}$. To create a margin, scale the x direction by $\frac{3}{4}$. The graphics matrix is diagonal but not 2 by 2. It is 3 by 3 to rescale a plane and 4 by 4 to rescale a space:

$$\text{Scaling matrices } S = \begin{bmatrix} \frac{3}{4} & & \\ & \frac{1}{2} & \\ & & 1 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & c_3 & \\ & & & 1 \end{bmatrix}.$$

That last matrix S rescales the x, y, z directions by positive numbers c_1, c_2, c_3 . The extra column in all these matrices leaves the extra 1 at the end of every vector.

Summary The scaling matrix S is the same size as the translation matrix T . They can be multiplied. To translate and then rescale, multiply vTS . To rescale and then translate, multiply vST . Are those different? Yes.

The point (x, y, z) in \mathbf{R}^3 has homogeneous coordinates $(x, y, z, 1)$ in \mathbf{P}^3 . This “projective space” is not the same as \mathbf{R}^4 . It is still three-dimensional. To achieve such a thing, (cx, cy, cz, c) is the same point as $(x, y, z, 1)$. Those points of projective space \mathbf{P}^3 are really lines through the origin in \mathbf{R}^4 .

Computer graphics uses **affine** transformations, *linear plus shift*. An affine transformation T is executed on \mathbf{P}^3 by a 4 by 4 matrix with a special fourth column:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} = \begin{bmatrix} T(1, 0, 0) & 0 \\ T(0, 1, 0) & 0 \\ T(0, 0, 1) & 0 \\ T(0, 0, 0) & 1 \end{bmatrix}.$$

The usual 3 by 3 matrix tells us three outputs, this tells four. The usual outputs come from the inputs $(1, 0, 0)$ and $(0, 1, 0)$ and $(0, 0, 1)$. When the transformation is linear, three outputs reveal everything. When the transformation is affine, the matrix also contains the output from $(0, 0, 0)$. Then we know the shift.

3. Rotation A rotation in \mathbf{R}^2 or \mathbf{R}^3 is achieved by an orthogonal matrix Q . The determinant is $+1$. (With determinant -1 we get an extra reflection through a mirror.) Include the extra column when you use homogeneous coordinates!

$$\text{Plane rotation } Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{becomes} \quad R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This matrix rotates the plane around the origin. **How would we rotate around a different point** (4, 5)? The answer brings out the beauty of homogeneous coordinates. **Translate** (4, 5) to (0, 0), **then rotate by θ , then translate** (0, 0) **back to** (4, 5):

$$v T_- R T_+ = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & -5 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 5 & 1 \end{bmatrix}.$$

I won't multiply. The point is to apply the matrices one at a time: v translates to $v T_-$, then rotates to $v T_- R$, and translates back to $v T_- R T_+$. Because each point $\begin{bmatrix} x & y & 1 \end{bmatrix}$ is a row vector, T_- acts first. The center of rotation (4, 5)—otherwise known as (4, 5, 1)—moves first to (0, 0, 1). Rotation doesn't change it. Then T_+ moves it back to (4, 5, 1). All as it should be. The point (4, 6, 1) moves to (0, 1, 1), then turns by θ and moves back.

In three dimensions, every rotation Q turns around an axis. The axis doesn't move—it is a line of eigenvectors with $\lambda = 1$. Suppose the axis is in the z direction. The 1 in Q is to leave the z axis alone, the extra 1 in R is to leave the origin alone:

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} Q & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now suppose the rotation is around the unit vector $\mathbf{a} = (a_1, a_2, a_3)$. With this axis \mathbf{a} , the rotation matrix Q which fits into R has three parts:

$$Q = (\cos \theta)I + (1 - \cos \theta) \begin{bmatrix} a_1^2 & a_1 a_2 & a_1 a_3 \\ a_1 a_2 & a_2^2 & a_2 a_3 \\ a_1 a_3 & a_2 a_3 & a_3^2 \end{bmatrix} - \sin \theta \begin{bmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{bmatrix}. \quad (1)$$

The axis doesn't move because $\mathbf{a} Q = \mathbf{a}$. When $\mathbf{a} = (0, 0, 1)$ is in the z direction, this Q becomes the previous Q —for rotation around the z axis.

The linear transformation Q always goes in the upper left block of R . Below it we see zeros, because rotation leaves the origin in place. When those are not zeros, the transformation is affine and the origin moves.

4. Projection In a linear algebra course, most planes go through the origin. In real life, most don't. A plane through the origin is a vector space. The other planes are affine spaces, sometimes called “flats.” An affine space is what comes from translating a vector space.

We want to project three-dimensional vectors onto planes. Start with a plane through the origin, whose unit normal vector is \mathbf{n} . (We will keep \mathbf{n} as a column vector.) The vectors in the plane satisfy $\mathbf{n}^\top \mathbf{v} = 0$. **The usual projection onto the plane is the matrix $I - \mathbf{n}\mathbf{n}^\top$.** To project a vector, multiply by this matrix. The vector \mathbf{n} is projected to zero, and the in-plane vectors \mathbf{v} are projected onto themselves:

$$(I - \mathbf{n}\mathbf{n}^\top)\mathbf{n} = \mathbf{n} - \mathbf{n}(\mathbf{n}^\top \mathbf{n}) = \mathbf{0} \quad \text{and} \quad (I - \mathbf{n}\mathbf{n}^\top)\mathbf{v} = \mathbf{v} - \mathbf{n}(\mathbf{n}^\top \mathbf{v}) = \mathbf{v}.$$

In homogeneous coordinates the projection matrix becomes 4 by 4 (but the origin doesn't move):

$$\text{Projection onto the plane } \mathbf{n}^T \mathbf{v} = 0 \quad P = \begin{bmatrix} I - \mathbf{n}\mathbf{n}^T & 0 \\ 0 & 1 \end{bmatrix},$$

Now project onto a plane $\mathbf{n}^T(\mathbf{v} - \mathbf{v}_0) = 0$ that does *not* go through the origin. One point on the plane is \mathbf{v}_0 . This is an affine space (or a *flat*). It is like the solutions to $A\mathbf{v} = \mathbf{b}$ when the right side is not zero. One particular solution \mathbf{v}_0 is added to the nullspace—to produce a flat.

The projection onto the flat has three steps. Translate \mathbf{v}_0 to the origin by T_- . Project along the \mathbf{n} direction, and translate back along the row vector \mathbf{v}_0 :

$$\text{Projection onto a flat} \quad T_- P T_+ = \begin{bmatrix} I & 0 \\ -\mathbf{v}_0 & 1 \end{bmatrix} \begin{bmatrix} I - \mathbf{n}\mathbf{n}^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ \mathbf{v}_0 & 1 \end{bmatrix}.$$

I can't help noticing that T_- and T_+ are inverse matrices: translate and translate back. They are like the elementary matrices of Chapter 2.

The exercises will include reflection matrices, also known as *mirror matrices*. These are the fifth type needed in computer graphics. A reflection moves each point twice as far as a projection—*the reflection goes through the plane and out the other side*. So change the projection $I - \mathbf{n}\mathbf{n}^T$ to $I - 2\mathbf{n}\mathbf{n}^T$ for a mirror matrix.

The matrix P gave a “parallel” projection. All points move parallel to \mathbf{n} , until they reach the plane. The other choice in computer graphics is a “perspective” projection. This is more popular because it includes foreshortening. With perspective, an object looks larger as it moves closer. Instead of staying parallel to \mathbf{n} (and parallel to each other), the lines of projection come *toward the eye*—the center of projection. This is how we perceive depth in a two-dimensional photograph.

The basic problem of computer graphics starts with a scene and a viewing position. Ideally, the image on the screen is what the viewer would see. The simplest image assigns just one bit to every small picture element—called a *pixel*. It is light or dark. This gives a black and white picture with no shading. You would not approve. In practice, we assign shading levels between 0 and 2^8 for three colors like red, green, and blue. That means $8 \times 3 = 24$ bits for each pixel. Multiply by the number of pixels, and a lot of memory is needed!

Physically, a *raster frame buffer* directs the electron beam. It scans like a television set. The quality is controlled by the number of pixels and the number of bits per pixel. In this area, one standard text is *Computer Graphics: Principles and Practices* by Foley, Van Dam, Feiner, and Hughes (Addison-Wesley, 1995). The newer books still use homogeneous coordinates to handle translations. My best references were notes by Ronald Goldman and by Tony DeRose.