9.6 An Application of Norms: Solving Inconsistent Linear Systems

The problem of solving an inconsistent linear system $Ax = b$ often arises in practice. This is a system where b does not belong to the column space of A, usually with more equations than variables. Thus, such a system has no solution. Yet we would still like to "solve" such a system, at least approximately.

Such systems often arise when trying to fit some data. For example, we may have a set of 3D data points

$$\{p_1, \ldots, p_n\},$$

and we have reason to believe that these points are nearly coplanar. We would like to find a plane that best fits our data points. Recall that the equation of a plane is

$$\alpha x + \beta y + \gamma z + \delta = 0,$$

with $(\alpha, \beta, \gamma) \neq (0, 0, 0)$. Thus, every plane is either not parallel to the x-axis ($\alpha \neq 0$) or not parallel to the y-axis ($\beta \neq 0$) or not parallel to the z-axis ($\gamma \neq 0$).

Say we have reasons to believe that the plane we are looking for is not parallel to the z-axis. If we are wrong, in the least squares solution, one of the coefficients, α, β, will be very large. If $\gamma \neq 0$, then we may assume that our plane is given by an equation of the form

$$z = ax + by + d,$$

and we would like this equation to be satisfied for all the p_i's, which leads to a system of n equations in 3 unknowns a, b, d, with $p_i = (x_i, y_i, z_i)$;

$$\begin{aligned} ax_1 + by_1 + d &= z_1 \\ \vdots \quad & \quad \vdots \\ ax_n + by_n + d &= z_n. \end{aligned}$$

However, if n is larger than 3, such a system generally has *no solution*. Since the above system can't be solved exactly, we can try to find a solution (a, b, d) that *minimizes the least-squares error*

$$\sum_{i=1}^{n} (ax_i + by_i + d - z_i)^2.$$

This is what Legendre and Gauss figured out in the early 1800's!

In general, given a linear system

$$Ax = b,$$

we solve the *least squares problem*: minimize $\|Ax - b\|_2^2$.