2.
a)

basis $C(A^T) = \{(\frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5})\}$
basis $N(A) = \{(\frac{2\sqrt{5}}{5}, -\frac{\sqrt{5}}{5})\}$
basis $C(A) = \{(\frac{\sqrt{10}}{10}, \frac{3\sqrt{10}}{10})\}$
basis $N(A^T) = \{(-\frac{3\sqrt{10}}{10}, \frac{\sqrt{10}}{10})\}$

3.

$Q = UV^T$

$= \begin{bmatrix} \frac{\sqrt{10}}{10} & -\frac{3\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} & \frac{\sqrt{10}}{10} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix}$

$= \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$

$H = V\Sigma V^T$

$= \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix}$

$= \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 2\sqrt{2} & 4\sqrt{2} \end{bmatrix}$

4.

$A^+ = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{10} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10} \\ -\frac{3\sqrt{10}}{10} & \frac{\sqrt{10}}{10} \end{bmatrix} = \begin{bmatrix} \frac{1}{50} & \frac{3}{50} \\ \frac{1}{25} & \frac{3}{25} \end{bmatrix}$

$P_{row} = A^+A = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$

$P_{col} = AA^+ = \begin{bmatrix} \frac{1}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{bmatrix}$

10.

$A = [3 \quad 4 \quad 0]$

$A^TA = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} [3 \quad 4 \quad 0]$

$= \begin{bmatrix} 9 & 12 & 0 \\ 12 & 16 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\lambda = 0, 0, 25$
$v_1 = \frac{\sqrt{5}}{5}(-\frac{4}{3}, 1, 0)$
$v_2 = (0, 0, 1)$
$v_3 = \frac{4}{5}(\frac{3}{4}, 1, 0)$

15.

$A^TA\hat{x} = A^Tb$

$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \hat{x} = \begin{bmatrix} \frac{1}{5} & \frac{3}{10} \\ \frac{1}{5} & \frac{1}{10} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \hat{x} = \begin{bmatrix} \frac{1}{5}b_1 + \frac{1}{10}b_2 \\ 0 \quad 0 \end{bmatrix}$

A^TA is singular

17.

$AA^+p = p$
$AA^+e = 0$
$A^+Ax_r = x_r$
$A^+Ax_n = 0$

18.

$A = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$A^TA = [3 \quad 4]\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$= [25]$

$\lambda = 25$

Case $\lambda = 25$

$(A - 25I)x_1 = 0$

$x_1 = (1)$

$v_1 = (1)$

$Av_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}[1]$

$= \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$\sigma_1 = 5, \; u_1 = (\frac{3}{5}, \frac{4}{5})$

choose $u_2 = (\frac{4}{5}, -\frac{3}{5})$

$A = U\Sigma V^T \Rightarrow \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} 5 \\ 0 \end{bmatrix} [1]$

$A^+ = V\Sigma^+U^T = [1][\frac{1}{5} \quad 0] \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}$

$= [\frac{3}{25} \quad \frac{4}{25}]$

$A^+A = [1]$

$AA^+ = \begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix}$

$x^+ = A^+b$

$= [\frac{3}{25} \quad \frac{4}{25}] \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1$

$x^+ = A^+b$

$= [\frac{3}{25} \quad \frac{4}{25}] \begin{bmatrix} -4 \\ 3 \end{bmatrix}$

$= 0$

Chapter 8: Applications

Chapter 8.1: Matrices in Engineering

- Engineering problems produce symmetric matrices K, which is often positive definite. This because a) $K = A^TA$ and $K = A^TCA$, and b) the expression $\frac{1}{2}u^TKu$ represents energy, which is never negative. The matrix C, often diagonal, contains positive physical constants
- Our first example: K is the stiffness matrix, and $K^{-1}f$ is the structure's responses to forces f from outside
- Engineering leads to linear algebra in two ways:
  - Direct way: the laws connecting position or velocity are linear. Expressed as matrix eqs.
  - Indirect way: the physical system is continuous; expressed as differential equations. To find accurate solutions, we approximate by finite difference or finite element equations

- We use the equilibrium equations $Ku = f$. With motion $M\frac{d^2u}{dt^2} + Ku = f$ becomes dynamic. Then we use eigenvalues from $Kx = \lambda Mx$, or finite differences.
- The matrices

$$K_0 = A_0^T A_0 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \qquad A_0^T C_0 A_0 = \begin{bmatrix} c_1+c_2 & -c_2 & 0 \\ -c_2 & c_2+c_3 & -c_3 \\ 0 & -c_3 & c_3+c_4 \end{bmatrix}$$

Fixed-fixed — Spring constants included

$$K_1 = A_1^T A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \qquad A_1^T C_1 A_1 = \begin{bmatrix} c_1+c_2 & -c_2 & 0 \\ -c_2 & c_2+c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{bmatrix}$$

Fixed-free — Spring constants included

$$K_{singular} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \qquad K_{circular} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Free-free

- The matrices $K_0, K_1, K_{singular}$, and $K_{circular}$ have $C = I$ for simplicity. This means all spring constants are $c_i = 1$

A Line of Springs

- Fig 8.1.1 shows 3 masses m_1, m_2, m_3 connected by a line of springs. One case has 4 springs, with top and bottom fixed. The fixed-free case has 3 springs: the lowest mass hangs freely. The fixed-fixed problem will lead to K_0 and $A_0^T C_0 A_0$. The fixed-free problem will lead to K_1 and $A_1^T C_1 A_1$. A free-free problem produces $K_{singular}$.
- We want equations for the mass movements u and the tensions y:
  $u = (u_1, u_2, u_3)$ = movement of the masses
  $y = (y_1, y_2, y_3, y_4)$ or (y_1, y_2, y_3) = tensions in the springs
- When a mass moves downward, its displacement is positive ($u_i > 0$). For the springs, tension is positive. Hooke's Law $y = ce$.
- We want to link these equations into $Ku = f$. f comes from gravity, so we get $f = (m_1 g, m_2 g, m_3 g)$

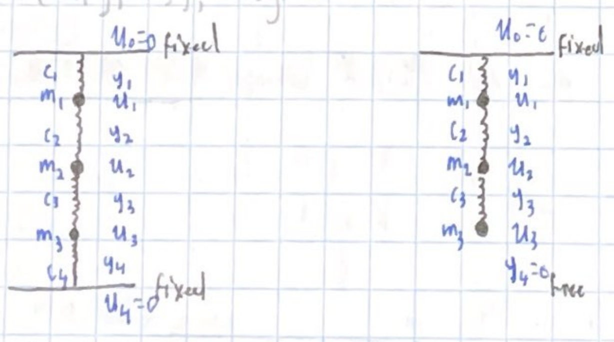


Fig 8.1.1: Fixed-fixed and fixed-free spring lines.

- We have the following quantities:

$u = (u_1, u_2, \dots, u_n)$ Movements of n masses — $y = (y_1, y_2, \dots, y_m)$ Internal forces in m springs

$e = (e_1, e_2, \dots, e_m)$ Elongation of m springs — $f = (f_1, f_2, \dots, f_n)$ External forces on n masses

- The framework that connects u to e to y to f looks like this:

$$\begin{tikzcd} \boxed{u} \arrow[d, "A"'] & \boxed{f} \\ \boxed{e} \arrow[r, "C"] & \boxed{y} \arrow[u, "A^T"'] \end{tikzcd}$$

$e = Au$	A is m by n
$y = Ce$	C is m by m
$f = A^T y$	A^T is n by m

- The elongations e is the displacements of each spring. When the system is vertical, the masses move by distances u_1, u_2, u_3. Each spring is displaced by $e_i = u_i - u_{i-1}$, the difference in displac

Stretching of each spring
$$\begin{aligned} e_1 &= u_1 && (u_0 = 0) \\ e_2 &= u_2 - u_1 \\ e_3 &= u_3 - u_2 \\ e_4 &= -u_3 && (u_4 = 0) \end{aligned}$$

- This is a 4 by 3 (m by n) difference matrix

$$e = Au \text{ is } \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -1 & 1 & \\ & -1 & 1 \\ & & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

- The next equation $y = Ce$ is just Hooke's law

$$\begin{aligned} y_1 &= c_1 e_1 \\ y_2 &= c_2 e_2 \\ y_3 &= c_3 e_3 \\ y_4 &= c_4 e_4 \end{aligned} \quad \text{is} \quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & c_3 & \\ & & & c_4 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

- Combining $e = Au$ with $y = Ce$, we get $y = CAu$
- The internal spring forces balance the gravitational forces
  - Each mass is pushed or pulled by the spring force above it (y_j). From below it feels the spring force y_{j+1} and f_j from gravity. Thus $y_j = y_{j+1} + f_j$ or $f_j = y_j - y_{j+1}$

$$\begin{aligned} f_1 &= y_1 - y_2 \\ f_2 &= y_2 - y_3 \\ f_3 &= y_3 - y_4 \end{aligned} \quad \text{is} \quad \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

  - This is $f = A^T y$. These combine into $Ku = f$, where the stiffness matrix is $K = A^T C A$

$$\left.\begin{aligned} e &= Au \\ y &= Ce \\ f &= A^T y \end{aligned}\right\} \text{ combine into } A^T C A u = f \text{ or } Ku = f$$

- In the language of elasticity, $e = Au$ is the kinematic equation (for displacement). The force balance $f = A^T y$ is the static equation (for equilibrium). The constitutive law is $y = Ce$

- With $C = I$, we get

$$K_0 = A_0^T A_0 = \begin{bmatrix} 2 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix}$$

- Properties:
  1. K is <u>tridiagonal</u>, because mass 3 is not connected to mass 1.
  2. K is <u>symmetric</u> because C is symmetric and A^T comes with A.
  3. K is <u>positive semidefinite</u> because $c_i > 0$ and A has independent columns
  4. K^{-1} is a full matrix with <u>all positive entries</u>
- Example
  - Suppose $c_i = c$ and $m_j = m$. Find the movements u and tensions y.

$$u = K^{-1} f = \frac{1}{4c}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} mg \\ mg \\ mg \end{bmatrix} = \frac{mg}{c}\begin{bmatrix} \frac{3}{2} \\ 2 \\ \frac{3}{2} \end{bmatrix}$$

  - The displacement u_2 is larger than the others

$$e = Au \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}\frac{mg}{c}\begin{bmatrix} \frac{3}{2} \\ 2 \\ \frac{3}{2} \end{bmatrix} = \frac{mg}{c}\left.\begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{3}{2} \end{bmatrix}\right\} \text{add to } 0$$

  - To find the spring force, multiply e by C: $y = (\frac{3}{2}mg, \frac{1}{2}mg, -\frac{1}{2}mg, -\frac{3}{2}mg)$
- Warning: Normally you cannot write $K^{-1} = A^{-1} C^{-1} (A^T)^{-1}$ bc A is rectangular

Fixed End and Free End

- Remove the fourth spring. All matrices become 3 by 3; the pattern does not change.

$$A_1^T C_1 A_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} c_1 & & \\ & c_2 & \\ & & c_3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

and (set $c_4 = 0$)

$$K_1 = A^T C_1 A_1 = \begin{bmatrix} c_1 + c_2 & -c_2 & 0 \\ -c_2 & c_2 + c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{bmatrix}$$

- Example
  - All $c_i = c$ and all $m_j = m$

$$K_1 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \text{ and } K_1^{-1} = \frac{1}{c}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

$$u = K_1^{-1} f = \frac{1}{c}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} mg \\ mg \\ mg \end{bmatrix} = \frac{mg}{c}\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$$

$3+2$

$3+2+1$

$$e = Au = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\frac{mg}{c}\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix} = \frac{mg}{c}\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Two Free Ends: K is Singular

- Now the matrix A is 2 by 3:

$$e = Au \Rightarrow \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} u_2 - u_1 \\ u_3 - u_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

- Now there is a nonzero to $Au = 0$. The masses can move without stretching the springs. The whole line can shift by $u = (1,1,1)$ and the leaves $u = (0,0)$. A has dependent columns and the vector $(1,1,1)$ is in its nullspace:

$$Au = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \text{no stretching}$$

- $Au = 0$ leads $A^TCAu = 0$, so A^TCA is only positive semidefinite, without C_1 and C_4. The pivots will be C_2 and C_3 and no third pivot. The rank is only 2

$$\begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C_2 & 0 \\ 0 & C_3 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} C_2 & -C_2 & 0 \\ -C_2 & C_2 + C_3 & -C_3 \\ 0 & -C_3 & C_3 \end{bmatrix}$$

- Two eigenvalues will be positive but $x = (1,1,1)$ is an eigenvector for $\lambda = 0$. We can solve $A^TCAu = f$ only for special vectors f. The forces need to add to $f_1 + f_2 + f_3 = 0$.

Circle of Springs

- A third spring will complete the circle from mass 3 to mass 1. K is still singular

$$K_{circular} = A^T_{circular} A_{circular} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

- $K_{circular}$ is always symmetric and semidefinite
- We have pivots 2 and $\frac{3}{2}$ and eigenvalues are 3 and 3 and 0. The nullspace still contains $x = (1,1,1)$

Continuous Instead of Discrete

- Matrix equations are discrete. Differential equations are continuous. We'll see the differential equations that corresponds to the tridiagonal $-1, 2, -1$ matrix A^TA.
- The matrices A and A^T correspond to the derivatives $\frac{d}{dx}$ and $-\frac{d}{dx}$. Remember that $e = Au$ takes differences $u_i - u_{i-1}$ and $f = A^Ty$ takes differences $y_i - y_{i+1}$

$$\frac{u_i - u_{i-1}}{\Delta x} \text{ is like } \frac{du}{dx} \qquad \frac{y_i - y_{i+1}}{\Delta x} = -\frac{dy}{dx}$$

- Δx didn't appear earlier, we imagine the distance between masses was 1.

$$e(x) = Au = \frac{du}{dx} \quad y(x) = c(x)e(x) \quad A^Ty = -\frac{dy}{dx} = f(x)$$

- Combining the equations, we get $A^TCAu(x) = f(x)$, we have a differential equation.
- The line of springs becomes an elastic bar

$$A^TCAu(x) = f(x) \text{ is } -\frac{d}{dx}\left(c(x)\frac{du}{dx}\right) = f(x)$$

- A^TA corresponds to a second derivative. A is a difference matrix and A^TA is a second derivative matrix. The matrix has $-1, 2, -1$ and the equation has $-d^2u/dx^2$

$-u_{i+1} + 2u_i - u_{i-1}$ is a second difference $\quad -\frac{d^2u}{dx^2}$ is a second derivative

- We usually have 3 choices for a derivative (forward, backward, centered difference)

$$\frac{du}{dx} \approx \frac{u(x+\Delta x) - u(x)}{\Delta x} \quad \text{or} \quad \frac{u(x) - u(x-\Delta x)}{\Delta x} \quad \text{or} \quad \frac{u(x+\Delta x) - u(x-\Delta x)}{2\Delta x}$$

- For $\frac{d^2u}{dx^2}$, we have $[u(x+\Delta x) - 2u(x) + u(x-\Delta x)]/(\Delta x)^2$. The signs are $-1, 2, -1$ because first derivative is antisymmetric, so the second differences are negative definite and we change to $-d^2u/dx^2$
- In scientific computing, we typically create the discrete matrix K by approximately the continuous problem

Fixed
Fixed
$$Au = \frac{1}{\Delta x}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \approx \frac{du}{dx} \quad \text{with } u_0 = u_4 = 0$$

Fixed
Free
$$A^Ty = \frac{1}{\Delta x}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \approx -\frac{dy}{dx} \quad \text{with } u_0 = y_4 = 0$$

- Process
  1. Model the problem with a differential equation
  2. Discretize the differential equation to a difference equation
  3. Understand and solve the difference equation (and boundary conditions)
  4. Interpret the solution; visualize it; redesign if needed

Problem Set 8.1

1.

$$\det\begin{bmatrix} c_1+c_2 & -c_2 & 0 \\ -c_2 & c_2+c_3 & -c_3 \\ 0 & -c_3 & c_3+c_4 \end{bmatrix} = (c_1+c_2)(c_2+c_3)(c_3+c_4) - (-c_2)(-c_2)(c_3+c_4) - (-c_3)(-c_3)(c_1+c_2)$$
$$= (c_1c_2 + c_1c_3 + c_2^2 + c_2c_3)(c_3+c_4) - c_2^2c_3 - c_2^2c_4 - c_1c_3^2 - c_2c_3^2$$
$$= c_1c_2c_3 + c_1c_3^2 + c_2^2c_3 + c_2c_3^2 + c_1c_2c_4 + c_1c_3c_4 + c_2^2c_4 + c_2c_3c_4 - c_2^2c_3 - c_2^2c_4 - c_1c_3^2 - c_2c_3^2$$
$$= c_1c_2c_3 + c_1c_2c_4 + c_1c_3c_4 + c_2c_3c_4$$

4.

$$-\frac{d}{dx}\left(c(x)\frac{du}{dx}\right) = f(x)$$

$c(x) \neq 0$

$$-c(x)\frac{du}{dx} = \int_{c_{end}}^{x} f(t)\,dt$$

this is zero at both endpoints

$$\int_{c_{end}}^{c_{end}} f(x)\,dx = 0$$

$$\int_{c_{end}}^{c_{start}} f(x)\,dx = 0$$

5.

$$\frac{dy}{dx} = f(x) \qquad y = \int_1^x f(t)\,dt$$

$$-dy = f(x)\,dx \qquad y = \int_1^x dt = [t]_1^x$$

$$-\int dy = \int f(x)\,dx \qquad = x - 1$$

$$-y + c = \int f(x)\,dx$$

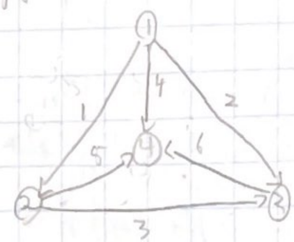
$$y = c - \int f(x)\,dx$$

$$y|_{x=1} = 0 = c - \int_1^1 f(t)\,dt \quad c = 0$$

Chapter 8.2: Graphs and Networks

- Graphs are entities of nodes connected by edges
- This section focuses on incidence matrices of graphs, which tell us how the n nodes are connected by the m edges. Normally $m>n$.
- For any m by n matrix, there are two fundamental subspaces in $\mathbb{R}^n$ and two in $\mathbb{R}^m$
  - We review the four subspaces for any matrix and constructed a directed graph and its incidence matrix
  - By specializing to incidence matrices, the laws of linear algebra become Kirchhoff's Laws
- Every entry in an incidence matrix is -1, 0, or 1. This holds during elimination. All pivots and multipliers are ± 1. Therefore both factors in $A=LU$ also had 0 or ± 1.
- Example: Differences in voltage across 6 edges of a graph. The columns are voltages of each node

$$A=\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \text{ reduces to } U=\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$



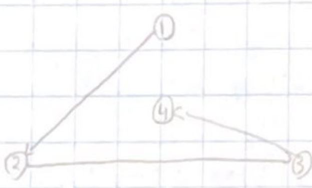
- The nullspace of A and U is line through $x=(1,1,1,1)$. The column space of A and U have dimension $r=3$. The pivot rows are a basis for the row space. $x=(1,1,1,1)$ in the nullspace is perpendicular to all the basis rows.
  - Equal voltages produce no current.

Directed Graphs and Incidence Matrices

- Looking at the graph row by row, we can establish the directed edges: $-1=$ source, $1=$ destination

$$\begin{array}{c|c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} & \text{node numbers} \\ A=\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} \text{ edge numbers} \end{array}$$

- Complete graph: every pair of nodes is connected by an edge
- Max edge: $\frac{1}{2}n(n-1)$



$$B=\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

- Tree: no closed loops
- min edges: $m=n-1$

- The rows of B match the non zero rows of U. Elimination reduces every graph to a tree.
- The loops produce zero rows in U

$$\text{edges } 1,2,3 \left\{ \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix} \right. \longrightarrow \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- These steps are typical. When two edges share a node, elimination gives the shortcut edge that skips the shared node.
- A idea: Rows are dependent when edges form a loop. Independent rows come from trees. This is the key for the row space.
- For the column space, Ax is a vector of differences

$$Ax = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_1 \\ x_3 - x_2 \\ x_4 - x_1 \\ x_4 - x_2 \\ x_4 - x_3 \end{bmatrix}$$

  - The unknowns represent potentials/voltage.
  - The vector Ax is then a vector of potential or electric differences

- The nullspace contains all solution to $Ax=0$ — all six potentials differences are 0, So all 4 potentials are the same. We can raise or increase all potentials by an arbitrary constant (like Calculus $+C$) without changing the differences or. If we set $x_4 = 0$ ("grounding it"), we can get absolute values without a constant
- An arbitrary vector is in the row space if it is orthogonal to $(1,1,1,1)$
- How can we tell if a particular b is in the column space of an incidence matrix
  - Ax is a vector of differences. Adding differences across a closed loop must sum to 0.
  - The components of Ax add to zero around every loop. When b is in the column space, its components in each loop sum to 0.
  - This is Kirchhoff's loop Law for voltages:
    The directed sum of the potential differences around any closed loop is 0
- Looking at the left nullspace:

$$A^Ty = \begin{bmatrix} -1 & -1 & 0 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

  - we have 3 equations ($r=3$)
  - The first equation is $-y_1 - y_2 - y_4 = 0$. The net flow into node 1 is 0.
  - The fourth equation says $y_4 + y_5 + y_6 = 0$. Flow into the node - Flow out is 0.
  - $A^Ty = 0$ is Kirchhoff's Current Law:
    Flow in equals flow out at each node
- What's the solution to $A^Ty = 0$, the currents must balance each other. The easiest way is to flow around a loop: If a unit of current goes around the big triangle (forward on edge 1 and 3, backward on edge 3), we get $y = (1,-1,1,0,0,0)$. This satisfies $A^Ty = 0$. Every loop current is a solution to the current law.
- We get 3 basis solutions, one for each small loop. The big loop is the sum of these basis vectors
- For every graph in a plane, linear algebra gives Euler's formula:
  (number of nodes) − (number of edges) + (number of small loops) = 1

Networks and A^TCA

- In a real network, the current y along an edge is the product of potential difference (voltage) and conductance c. c is decided by the material.
- The graph is known from its "connectivity matrix", which tells us the connections between nodes and edges. A network goes further and assigns a conductance c_i to each. These constants go into the diagonal matrix C.
- For a network of resistors, the conductance c is $1/(\text{resistance})$
  - From this we have ohms law: $I = \frac{V}{R}$
- Ohm's Law for all m currents is $y = -CAx$. Ax gives potential differences and C multiplies by conductance. Combining with Kirchhoff's Law, we get $A^TCAx = 0$. 0 shouldn't be on the right side. We want external power. Ex. $A^TCAx = (5, 0, 0)$ 5 volts from node 1.
- In circuit theory, we change Ax to $-Av$. The flow is from higher to lower potential.

Chapter 8.3: Markov Matrices, Population, and Economics

- This section is about positive matrices: every $a_{ij} > 0$. The largest eigenvalue is real and positive and so is its eigenvector.
- In economics, ecology, and population dynamics, this facts leads a long way.

  Markov $\lambda_{max} = 1$ population $\lambda_{max} > 1$ consumption $\lambda_{max} < 1$

  - λ_{max} controls the powers of A

Markov Matrices

- Suppose we multiply a vector $u_0 = (a, 1-a)$ again and again by this A

  Markov Matrix $A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$ $u_1 = Au_0$ $u_2 = A^2u_0$ $\ldots$ $u_n = A^n u_0$
- If we take the limit, we reach a steady state

  $\lim_{k \to \infty} A^k u_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$ regardless of the "a" we choose for u_0
  - Doing $\lim_{k \to \infty} Au_k = \lim_{k \to \infty} u_k$ gives that $\lim_{k \to \infty} u_k$ is an eigenvector with eigenvalue 1.
- A Markov Matrix:
  1. Every entry of A is non-negative
  2. Every column of A sums to 1.
- A Markov Matrix has the following properties
  1. Multiplying a nonnegative u_0 by A produces a nonnegative $u_1 = Au_0$
  2. If the components of u_0 sum to 1, so do the components of $u_1 = Au_0$
     - The components of u_0 sum to 1 when $[1 \ldots 1] u_0 = 1$. This is true for each column of A by property 2. Then by matrix multiplication $[1 \ldots 1] A = [1 \ldots 1]$

       components of Au_0 add to 1 $[1 \ldots 1] Au_0 = [1 \ldots 1] u_0 = 1$
- Every vector $A^k u_0$ is nonnegative with components that sum to 1.

- Example 1
  - The fraction of rental cars in Denver starts at $\frac{1}{50} = 0.02$. The fraction outside is $\frac{49}{50} = 0.98$. Every month, 80% of the Denver cars stay in Denver, and 20% leave. Also 5% of outside cars come in (95% stay out). Then we get the Markov Matrix

$$A = \begin{bmatrix} .80 & .05 \\ .20 & .95 \end{bmatrix} \begin{matrix} \leftarrow \text{cars in} \\ \leftarrow \text{cars out} \end{matrix} \quad \text{leads to } u_1 = Au_0 = A\begin{bmatrix} .02 \\ .98 \end{bmatrix} = \begin{bmatrix} .065 \\ .935 \end{bmatrix}$$

  - Notice that $0.065 + 0.935 = 1$
  - We calculate powers with the factorization $A = S\Lambda S^{-1} \Rightarrow A^k = S\Lambda^k S^{-1}$
  - Diagonalizing, we get $\lambda = 1$ and $.75$ and $x = (.2, .8)$ and $x = (-1, 1)$
    We write u_0 as a combination of eigenvectors

$$u_0 = \begin{bmatrix} .02 \\ .98 \end{bmatrix} = \begin{bmatrix} .2 \\ .8 \end{bmatrix} + .18\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$A^k u_0 = A^k\begin{bmatrix} .2 \\ .8 \end{bmatrix} + .18A^k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$= \underbrace{(1)^k}_{\text{steady state}}\begin{bmatrix} .2 \\ .8 \end{bmatrix} + .18\underbrace{(.75)^k}_{1<|\lambda| \Rightarrow \text{decay}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \lim_{k\to\infty} A^k u = (1)\begin{bmatrix} .2 \\ .8 \end{bmatrix} = \begin{bmatrix} .2 \\ .8 \end{bmatrix}$$

  - Eventually we get that 20% of cars are in Denver and 80% outside, no matter u_0, as long as its components sum to 1.
- If A is a positive markov matrix, then $\lambda_1 = 1$ is the largest eigenvalue and eigenvector x_1 is the steady state

$$u_k = x_1 + c_2(\lambda_2)^k x_2 + \dots + c_n(\lambda_n)^k x_n \text{ always approaches } u_\infty = x_1$$

- No eigenvalue can have $|\lambda| > 1$, but you should watch out that another eigenvalue has $|\lambda| = 1$
- Example
  - $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has no steady state because $\lambda_2 = -1$.
    This matrix sends all cars inside Denver outside and vice versa, so the states alternate without reaching a steady state
  - If A had all entries strictly positive, $\lambda = 1$ is strictly larger than any other eigenvalue

Perron-Frobenius Theorem
- Applies to matrices with $a_{ij} \geq 0$
- For the strict equality $a_{ij} > 0$, all numbers in $Ax = \lambda_{max} x$ are positive

Population Growth

- Divide the population into 3 groups: $\underset{n_1}{<20}$, $\underset{n_2}{20 \le age < 40}$, $\underset{n_3}{40 \le age < 60}$.
  Twenty years later, the sizes have changed for 2 reasons:
  1. Reproduction: $n_1^{new} = F_1 n_1 + F_2 n_2 + F_3 n_3$ gives a new generation
  2. Survival: $n_2^{new} = P_1 n_1$ and $n_3^{new} = P_2 n_2$ gives the older generations
- The fertility rates are F_1, F_2, F_3 (F_2 largest). The Leslie Matrix A might look like:

$$\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}^{new} = \begin{bmatrix} F_1 & F_2 & F_3 \\ P_1 & 0 & 0 \\ 0 & P_2 & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} .04 & 1.1 & .01 \\ .98 & 0 & 0 \\ 0 & .92 & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

  - This is population projection at its simplest, with a constant A. More advanced projections may have changing A's
- The matrix has $A \ge 0$ but not $A > 0$. The Perron-Frobenius theorem still applies because $A^3 > 0$. The largest eigenvalue is $\lambda_{max} \approx 1.06$. You can see the generations move

$$A^2 = \begin{bmatrix} 1.08 & 0.05 & .00 \\ 0.04 & 1.08 & .01 \\ 0.90 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0.10 & 1.19 & .01 \\ 0.06 & 0.05 & .00 \\ .04 & .99 & .01 \end{bmatrix}$$

- λ_{max} is the asymptotic growth rate

Linear Algebra in Economics: The Consumption Matrix

- The consumption matrix tells how much of each input goes into a unit of output. This describes the manufacturing side of the economy.
- We have n industries like chemicals, food, and oil. To produce a unit of chemicals, we may require .2 units of chemicals, .3 units of food and .4 units of oil. Those numbers go into row 1 of the consumption matrix.

$$\begin{bmatrix} \text{chemical output} \\ \text{food output} \\ \text{oil output} \end{bmatrix} = \begin{bmatrix} .2 & .3 & .4 \\ .4 & .4 & .1 \\ .5 & .1 & .3 \end{bmatrix} \begin{bmatrix} \text{chemical input} \\ \text{food input} \\ \text{oil input} \end{bmatrix}$$

- Similarly, row 2 shows the inputs to produce food and row 3 shows the inputs to produce oil.
- The question here is: Can the economy meet demands y_1, y_2, y_3 for chemicals, food, and oil. To do that, the inputs p_1, p_2, p_3 will have to be higher because part of p is consumed to produce y. The input is p and the consumption of Ap, which leaves the output $p - Ap$. This net production meets the demand y.
  - Find the vector p such that $p - Ap = y$ or $p = (I-A)^{-1}y$.
    So when is $(I-A)^{-1}$ a nonnegative matrix
- If A is small compared to I, then Ap is small compared to p. There is plenty of output. If A is too large, then production consumes more than it yields. In this case the external demand y cannot be met

- "Small" or "large" is decided by the largest eigenvalue λ_1 of A (which is positive)
  - If $\lambda_1 > 1$ then $(I-A)^{-1}$ has negative entries
  - If $\lambda_1 = 1$ then $(I-A)^{-1}$ fails to exist
  - If $\lambda_1 < 1$ then $(I-A)^{-1}$ is nonnegative as desired
- We have an infinite series $1 + x + x^2 + \dots = 1/(1-x)$ if x lies between -1 and 1. Similarly, for matrices, we have $(I-A)^{-1} = I + A^2 + A^3 + \dots$
  - If we multiply $S = I + A + A^2 + \dots$ by A, we get the series except for I. Therefore $S - AS = I \Rightarrow (I-A)S = I$. Therefore $S = (I-A)^{-1}$ if it converges. It converges if all λ of A have $|\lambda| < 1$
- Example:

$$A = \begin{bmatrix} .2 & .3 & .4 \\ .4 & .4 & .1 \\ .5 & .1 & .3 \end{bmatrix} \text{ has } \lambda_{max} = .9 \text{ and } (I-A)^{-1} = \frac{1}{93}\begin{bmatrix} 41 & 25 & 27 \\ 33 & 36 & 24 \\ 34 & 23 & 36 \end{bmatrix}$$

  - This economy is productive. A is small compared to I, because $\lambda_{max} = .9$. To meet the demand y, start from $p = (I-A)^{-1}y$. Then Ap is consumed in production, leaving $p - Ap$. This is $(I-A)p = y$ and the demand is met.
- Example

$$A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix} \text{ has } \lambda_{max} = 2 \text{ and } (I-A)^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}$$

  - This consumption matrix A is too large. Demands cannot be met, because the production consumes more than it yields

Problem Set 8.3

1.

$$A = \begin{bmatrix} .90 & .15 \\ .10 & .85 \end{bmatrix}$$

$$(A - I)x_1 = \begin{bmatrix} -.10 & .15 \\ .10 & -.15 \end{bmatrix} x_1 = 0$$

$$x_1 = (3, 2)$$

$$(A - 0.75I)x_2 = \begin{bmatrix} .15 & .15 \\ .10 & .10 \end{bmatrix} x_2 = 0$$

$$x_2 = (1, -1)$$

$$A = S\Lambda S^{-1} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & .75 \end{bmatrix}\begin{bmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{2}{5} & -\frac{3}{5} \end{bmatrix}$$

15.

$$(I-A)^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$(I-A)^{-1} = \begin{bmatrix} 1 & -4 \\ -.2 & 1 \end{bmatrix}^{-1} = \frac{1}{1.8}\begin{bmatrix} 1 & 4 \\ .2 & 1 \end{bmatrix}$$

Chapter 8.4: Linear Programming

- Linear programming = linear algebra + inequalities + minimization
- We start at $Ax=b$, but the only acceptable solutions are nonnegative.
  The matrix has $n \geq m$; more unknowns than equations. If there are solutions, there are infinite. Linear programming picks the solution $x^* \geq 0$ that minimizes cost.
  - The cost is $c_1x_1 + \dots + c_2x_n$
- Thus, a linear programming problem starts with a matrix A and two vectors b and c:
  i) A has $n > m$; for example, $A = [1 \ 1 \ 2]$
  ii) b has m components for m equations $Ax=b$; for example $b=[4]$
  iii) The cost vector c has n components; for example $c = [5 \ 3 \ 8]$
- We want the solution to $Ax=b$ that is $x \geq 0$ and minimizes $c \cdot x$
- Example:
  - Minimize $5x_1 + 3x_2 + 8x_3$ subject to $x_1 + x_2 + 2x_3 = 4$ and $x \geq 0$

$$Ax=b \Rightarrow \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [4], \quad c = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}$$

  - The equation $x_1 + x_2 + 2x_3 = 4$ gives a plane in three dimensions. The constraints $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$ chops the plane into a triangle. The solution x^* must lie in such triangle, illustrated in fig 8.4.1
  - Inside that triangle, all components of x are positive. On the edges of PQR, one component is zero. At the corners P, Q, R, two components are zero. The optimal solution x^* will be in one of these corners
  - These possible x's are the <u>feasible points</u>, and the triangle is the <u>feasible set</u>
  - The vectors that have zero cost are on the plane $5x_1 + 3x_2 + 8x_3 = 0$. This plane does not meet the triangle. So we increase the cost C until the plane $5x_1 + 3x_2 + 8x_3 = C$ meets the triangle

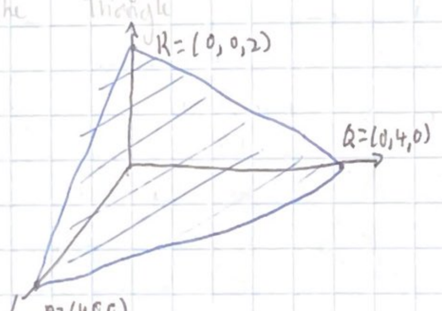


fig 8.4.1
The triangle contains all nonnegative solutions. The lowest cost solution is a corner P, Q, or R

- The first plane $5x_1 + 3x_2 + 8x_3 = C$ to touch the triangle has minimum cost C. The intersection point is the solution x^*. This point must be a corner P, Q, R. A moving plane cannot touch the interior without touching the corners first.
  - Check the cost of each corner

    $P = (4,0,0)$ costs 20 | $Q = (0,4,0)$ costs 12 | $R = (0,0,2)$ costs 16

  - Then $x^* = (0,4,0)$

- Note 1: Some linear programs maximize profit instead of minimizing cost. The mathematics are very similar. We start with a large value of C and decrease it until we have a solution.
- Note 2: It may happen that $Ax=b$ and $x \geq 0$ is impossible to satisfy. Then our feasible set is empty.
- Note 3: It may also occur that the feasible set is unbounded. The requirement $x_1+x_2-2x_3=4$ has P and Q as candidates, but R has moved to infinity
- Note 4: It is possible for the minimum cost to be $-\infty$ if we are dealing with an unbounded feasible set.

The Primal and Dual Problems

- Example:
  - The unknowns x_1, x_2, x_3 are hours of work for a Ph.D., a student, and a machine. The costs per hour are \$5, \$3, and \$8. The hours worked cannot be negative: $x \geq 0$
  - The Ph.D and student can do 1 problem per hour. The machine does 2 per hour. They collaborate on their homework, which has 4 problems $x_1+x_2+2x_3=4$. The cost is $5+3+8=16$ if all of them put in 1 hour. However, the Ph.D. should be put out of work by the student, who works as fast and costs less.
  - Setting x_1 to 0, we get a minimum cost of $2 \times 3+8=14$. However, the best cost is achieved by getting the student to do everything $4 \times 3=12$
- When $Ax=b$ has m equations, corners have m nonzeros. We solve $Ax=b$ for those m variables, with the other $n-m$ free variables set to zero. However, we don't know which m variables to choose. We need to choose all combinations. That is m choose n, or $n!/m!(n-m)!$. If we had $n=20$ and $m=8$, that's 5 billion combinations. Not good!
- The Dual Problem
  - In linear programming, problems come in pairs. The original problem and its dual. The original problem has A and two vectors b and c. The dual problem transposes A and swaps b and c. Maximize $b \cdot y$. Here's the dual to our example:
  - A cheater offers to solve homework problems by selling the answers. The charge is y dollars per problem, or $4y$ altogether ($b \cdot y$). The cheater must be as cheap as the Ph.D, student, and machine: $y \leq 5$, $y \leq 3$, $2y \leq 4$.
  - Maximize $b \cdot y$ subject to $A^Ty \leq c$. The maximum is $y=3$, maximizing profit at \$12, which matches the minimum in the original. This is the duality
- If either problem has a best vector x^* or y^*, then so does the other. Minimum cost $c \cdot x^*$ equals maximum income $b \cdot y^*$. This is the <u>strong duality theorem.</u>
- The cheater's income b^Ty cannot exceed the honest cost
  If $Ax=b$, $x \geq 0$, $A^Ty \leq c$ then $b^Ty=(Ax)^Ty=x^T(A^Ty) \leq x^Tc$
- The full theorem states that b^Ty at its max equals x^Tc at its minimum. The dot product of $x \geq 0$ and $s=c-A^Ty \geq 0$ gave $x^Ts \geq 0$. This is $x^TA^Ty \leq x^Tc$
- Equality needs $x^Ts=0$, so the optimal solution has $x_j^*=0$ or $s_j^*=0$ for each j

The Simplex Method

- The Simplex Method goes from one corner to a neighboring corner of lower cost
  - A corner is a vector $x \geq 0$ that satisfies the m equations $Ax=b$ with at most m positive components. The other $n-m$ components are 0. Those are the free variables, which we set to 0 to solve the m components. All the m component must be non-negative or it is a false corner.
  - A neighboring corner has a zero component of x become positive and one positive component become zero
  - The simplex method chooses which component "leaves" (becomes zero) and which "enters" (becomes positive) such that the total cost is lowered
- Here's the overall algorithm:
  - Look at every zero component of the current corner. If it changes from 0 to 1, the other nonzeroes need to adjust to keep $Ax=b$. Find the new x by back substitution and compute the change in cost $c \cdot x$. The change is the reduced cost "r" of the new component.
  - The entering variable is the one that gives the most negative r.
- Example
  - Suppose the current corner is $P=(4,0,0)$ with the Ph.D doing all the work. The current cost is \$20. Now suppose the student does 1 hour of work: $x=(3,1,0)$. Then $c \cdot x = \$18$. If the machine works one hour, we get $x=(3,0,1)$. Then $c \cdot x = \$18$. Both reduced costs are $r=-2$. The simplex method may choose either as the entering variable.
  - Now we move to selecting the leaving variable, the Ph.D, setting it to 0. Then we get $x=(0,4,0)$, a neighboring corner. We repeat again until all reduced costs are positive, meaning we are at the optimal corner, x^*.
- Generally, the Simplex Method takes $O(n)$ steps, but can take $O(2^n)$ in certain cases. A new approach is more complex but takes fewer steps.
- Example
  - Minimize the cost $c \cdot x = 3x_1 + x_2 + 9x_3 + x_4$ with $x \geq 0$ and

$$Ax=b \Rightarrow \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \qquad \begin{matrix} m=2 \text{ equations} \\ n=4 \text{ unknowns} \end{matrix}$$

  - Starting at corner $x=(4,2,0,0)$ which costs $c \cdot x = 14$. It has $m=2$ nonzeroes and $n-m=2$ zeroes. Try one unit each of them
    - $x_3=1$, $x_4=0$, then $x=(2,1,1,0)$ costs 16 $\quad r=2$
    - $x_4=1$, $x_3=0$, then $x=(3,3,0,1)$ costs 13 $\quad r=-1$ Entering
  - x_1 is leaving, giving us $x=(0,6,0,4)$. Trying more steps
    - $x_1=1$, $x_3=0$, then $x=(1,5,0,3)$ costs 11
    - $x_3=1$, $x_1=0$, then $x=(0,3,1,2)$ costs 14

    } both r's are positive $\Rightarrow x^*=(0,6,0,4)$

Interior Point Methods

- Interior point methods inside the feasible set, where $x > 0$, hoping to reach x^*
- To stay inside, we put a barrier at the boundary. Add an extra cost as a logarithm that blows up when any variable x_j touches 0.

The number θ is a small parameter that we move towards zero

Barrier Problem: Minimize $c^T x - \theta(\log x_1 + \dots + \log x_n)$ with $Ax = b$

- The cost is non-linear. The constraints $x_j \geq 0$ are not needed.
- The barrier gives an approximate problem for θ. The m constraints $Ax = b$ have Lagrange multipliers $y_1, \dots, y_m$. This is the good way to deal with constraints

y from Lagrange: $L(x, y, \theta) = c^T x - \theta(\sum \log x_i) - y^T(Ax - b)$

- $\partial L / \partial y = 0$ brings back $Ax = b$.

Optimality in barrier pblm $\frac{\partial L}{\partial x_j} = c_j - \frac{\theta}{x_j} - (A^T y)_j = 0$ which is $x_j s_j = \theta$

- The true problem is $x_j s_j = 0$. The barrier problem has $x_j s_j = \theta$. The solutions $x^*(\theta)$ lie on the central path to $x^*(0)$. We can solve the n optimality equations $x_j s_j = \theta$ can be solved iteratively via Newtons Method

Chapter 8.5: Fourier Series: Linear Algebra for Functions

- We will transition to infinite dimensional vector spaces
- What does it mean for a vector to have infinite components?
  - The vector becomes $v = (v_1, v_2, \dots)$
  - The vector becomes a function $f(x)$
- We will go both ways, and they will be connected by the idea of a Fourier series
- Dot products would be $v \cdot w = v_1 w_1 + v_2 w_2 + \dots$, an infinite series.
  - That poses the question, does $v \cdot w$ converge.
- We include all vectors $v = (v_1, v_2, v_3, \dots)$ in our infinite-dimensional Hilbert Space if and only if its length $\|v\|$ is finite:

$\|v\|^2 = v \cdot v = v_1^2 + v_2^2 + v_3^2 + \dots$ must add to a finite number

- Example:
  - The vector $v = (1, \frac{1}{2}, \frac{1}{4}, \dots)$ is included in the Hilbert Space. Its length is

$\sqrt{v_1^2 + v_2^2 + \dots} = \sqrt{\frac{1}{1 - \frac{1}{4}}} = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3}$

- If v and w have finite length, how large can their dot product be?
  - The sum $v \cdot w = v_1 w_1 + v_2 w_2 + \dots$ also adds to a finite number. The Swartz inequality still holds

Swartz inequality $|v \cdot w| \leq \|v\| \|w\|$

  - Even in infinite dimensional space the ratio of $v \cdot w$ to $\|v\|\|w\|$ is $\cos\theta$

- Now change to functions; The space of functions defined for $0 \le x \le 2\pi$: It must somehow be bigger than $\mathbb{R}^n$.
- We redefine the inner product and length as integrals

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,dx \quad \text{and} \quad \|f\|^2 = \int_0^{2\pi} [f(x)]^2\,dx$$

- Example:
  - $f(x) = \sin x$

$$\|f\|^2 = \langle f, f\rangle = \int_0^{2\pi} \sin^2 x\,dx = \left[\frac{1}{2}\sin x\cos x - \frac{1}{2}\int dx\right]_0^{2\pi} = \left(\tfrac{1}{2}\sin(2\pi)\cos(2\pi) - \tfrac{1}{2}(2\pi)\right) - \left(\tfrac{1}{2}\sin(0)\cos(0) - \tfrac{1}{2}(0)\right) = \pi$$

  Then, $\|f\| = \sqrt{\pi}$

  - Moreover, $\sin x$ and $\cos x$ are orthogonal

  Inner product is zero $\int_0^{2\pi} \sin x \cos x\,dx = \frac{1}{2}\int_0^{2\pi} \sin(2x)\,dx = \frac{1}{4}\left[\cos 2x\right]_0^{2\pi} = 0$

- Every function in the list is orthogonal to the others:
  $\cos(0x), \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \ldots$

Fourier Series

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

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

- We have an orthogonal basis! Our functions repeat every 2π.
- Remember, our list is infinite, We avoided $v = (1, 1, \ldots)$ because its infinite length, So we avoid functions like $\frac{1}{2} + \cos x + \cos 2x + \cos 3x$ (Note: This is π times the Dirac Delta function). All points inside $0 < x < 2\pi$ has a finite height. $\int \delta^2(x)\,dx$ is ∞, So it is not included in the Hilbert space
- Compute the length $f(x)$

$$(f, f) = \int_0^{2\pi} (a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + \ldots)^2\,dx$$

$$= \int_0^{2\pi} (a_0^2 + a_1^2\cos^2 x + b_1^2\sin^2 x + a_2^2\cos^2 2x + \ldots)\,dx$$

$$\|f\|^2 = 2\pi a_0^2 + \pi(a_1^2 + b_1^2 + a_2^2 + \ldots)$$

  - The step 1 to 2 uses orthogonality. All products like $\cos x \cos 2x$ integrate to zero. Line 2 contains what's left
  - If we divide each term by its length, we get an orthonormal basis for our function space

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos x}{\sqrt{\pi}}, \quad \frac{\sin x}{\sqrt{\pi}}, \quad \frac{\cos 2x}{\sqrt{\pi}}, \ldots$$

- We combine the basis with coefficients $A_0, A_1, B_1, A_2, \ldots$ to yield a function $F(x)$. Then the 2π and π's leave the formula

Function length = vector length $\|F\|^2 = \langle F, F\rangle = A_0^2 + A_1^2 + B_1^2 + A_2^2$

- The function has finite length exactly when the vector of coefficients has finite length.
- The Fourier Series connects our function space and infinite dimensional Hilbert Space. The function is in L^2, the coefficients in ℓ^2.
- The function space contains $f(x)$ exactly when the Hilbert space contains the vector $v = (a_0, a_1, b_1, \ldots)$ of Fourier coefficients. Both $f(x)$ and v have finite length
- Example
  - Suppose $f(x)$ is a square wave:

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < \pi \\ -1 & \text{if } \pi \le x < 2\pi \end{cases} \quad \text{forever}$$

  - This is an odd function, and so all terms in its Fourier Series are sines

$$f(x) = \frac{4}{\pi}\left[\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \ldots\right]$$

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

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

  - At $x = 0$, the sines are zero and so is the Fourier Series. This is halfway up the jump from -1 to $+1$. Also, at $x = \frac{\pi}{2}$, the square wave equals 1 and the sines alternate sign

Formula for π: $1 = \frac{4}{\pi}\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots\right) \Rightarrow \pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots\right)$

The Fourier Coefficients

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

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x$$

  - Multiply both sides by $\cos x$ and then integrate from 0 to 2π. From orthogonality, all integrals on the right side are 0 except $\cos^2 x$.

$$\int_0^{2\pi} f(x) \cos x \, dx = \int_0^{2\pi} a_1 \cos^2 x \, dx = \pi a_1$$

Divide by π to get a_1. We can do this for all a and b.

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

  - The integral of $f(x)\sin kx$ was $4/k$ for odd k in the square wave.
  - The constant term is excluded
  - The constant term a_0 is given by:

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

Compare Linear Algebra in $\mathbb{R}^n$

- Notice how similar the infinite and finite dimensional cases are. We can still use an orthogonal basis even in the infinite case.
- If we had a finite vector space, we construct construct a vector b using the orthogonal basis:

  $$b = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$$

  - Multiply both sides by v_1^T; orthogonality leads to zeroes

  coefficient c_1: $v_1^T b = c_1 v_1^T v_1 + 0 + 0 + \dots + 0$. Therefore $c_1 = v_1^T b / v_1^T v_1$

  $$c_n = \frac{\langle \text{nth basis vector}, b \rangle}{\|\text{nth basis vector}\|^2} = \frac{v_n^T b}{v_n^T v_n} = \frac{\int \cos kx \, f(x)\, dx}{\pi}$$, where $\langle \cdot , \cdot \rangle$ is the inner product

Chapter 8.6: Linear Algebra for Statistics and Probability

- Data tends to go in rectangular matrices so we'd expect a lot of A^TA
- This chapter will move to beyond least squares.

Weighted Least Squares

- In least squares, some measurements will be more reliable or important than others, and least squares must reflect that.
- To include weights in the m equations $Ax=b$, multiply each equation i by a weight w_i. We replace $Ax=b$ with $WAx=Wb$ and we use least squares just as before. $A^TA\hat{x} = A^Tb \Rightarrow (WA)^T(WA)\hat{x} = (WA)^TWb$. Let $C = W^TW$. Then:

  weighted least squares: $C = W^TW$ is in the n equations for $\hat{x}$: $A^TCA\hat{x} = A^TCb$

- 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 $\hat{x}_W = \frac{w_1^2 b_1 + \dots + w_m^2 b_m}{w_1^2 + \dots + w_m^2}$

- How do we choose the weights w_i? That 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 $\sigma_1, \dots, \sigma_m$, all variance will equal 1. So the weight is $w_i = 1/\sigma_i$ and the diagonal of $C = W^TW$ 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 they are dependent, off-diagonal entries show up in the covariance matrix Σ. The good choice is still $C = \Sigma^{-1}$.

Mean and Variance

- Two crucial numbers for a random variable are its mean and variance m and 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 from the mean)
- For a discrete random variable, the error e_i has probability p_i (the p_i add to 1)

  Mean $m = E[e] = \sum_1^n e_i p_i$ variable $\sigma^2 = E[(e-m)^2] = \sum_1^n (e_i - m)^2 p_i$

- Example
  - Flip a coin. The result is 0 or 1 (for heads or tails). These events both have $p_0 = p_1 = \frac{1}{2}$.

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

- Example:
  - Flip the 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
  - For all N, the number of ways to see M heads is N choose M. Divide that by the total 2^N possible outcomes to get the probability for each M.
  - The variance is $\sigma^2 = \frac{N}{4}$

The Covariance Matrix

- If we run m different experiments at once, they could be independent or there might be some correlation between them. Each measurement b is now a vector with n components. These components are the output 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, their product e_ie_j also has mean 0. However, if the measurements are at the same time by the same observer, e_i and e_j tend to have the same sign or size. The errors could be correlated. The products e_ie_j are weighted by p_{ij} (their probability): covariance $\sigma_{ij} = \sum_1 \sum_1 p_{ij}e_ie_j$. The sum of $e_i^2 p_{ii}$ is the variance σ_i^2.

  Covariance $\sigma_{ij} = \sigma_{ji} = E[e_ie_j] =$ expected value of e_ie_j

  This is the ijth and jith entry of the covariance matrix Σ. The iith entry is σ_i^2

Principle Component Analysis

- Start by measuring m properties of n samples (e.g. grades in m courses for n students). From each row, subtract its average so the sample means are zeros.

  We look for a combination of courses and/or a combination of students for which the data provides the most information.
- Information is "distance" from randomness and is measured by variance. A large variance in course grades means greater information than a small variance.
- The key matrix idea is the SVD $A = U\Sigma V^T$. The singular values in the diagonal matrix Σ 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
  - Suppose the grades A, B, C, F are worth 4, 2, 0, and -6 points. If each course and each student has one of each grade, then all means are 0

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

$\sigma = 12, 8, 4$ written as 3, 2, 1 to keep all entries as integers

- 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

Chapter 8.7: Computer Graphics

- Computer graphics deal with 4 transformations to move objects around in 3D space. These objects are projected onto 2D space to make images. These transformation matrices are 4 by 4 because they also must translate objects, which is not possible with 3D linear transformations.
- These transformations are:
  - 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 through P_0
  - Projection: onto a plane through the origin or P_0
- We change the origin's coordinates to $(0,0,0,1)$. The homogeneous coordinates of (x,y,z) are $(x,y,z,1)$
- 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 $\mathbb{R}^3$. Using homogeneous coordinates, the 4 by 4 matrix T shifts the whole space by v_0:

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

  - Computer graphics work with row vector times matrix instead of matrix times column vector
  - With the translation matrix, we have $[0\ 0\ 0\ 1]T = [x_0\ y_0\ z_0\ 1]$
  - To translate a row vector v, do $[v\ 1]T = [v_{new}\ 1]$
- Scaling
  - To scale, we multiply the homogeneous coordinate by S

Scaling Matrix $S = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

  - If instead, we have cI (4 by 4), multiplying a point gives (cx, cy, cz, c), which represents the same point as $(x,y,z,1)$. The special property of homogeneous coordinates is that multiplying cI does not move the point
  - We can scale by different factors in each direction with

Scaling Matrix $S = \begin{bmatrix} c_x & 0 & 0 & 0 \\ 0 & c_y & 0 & 0 \\ 0 & 0 & c_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

- The point (x,y,z) in $\mathbb{R}^3$ has homogeneous coordinates $(x,y,z,1)$ in $\mathbb{P}^3$. This "projective space" is not the same as $\mathbb{R}^4$. It is still three-dimensional. To achieve such a thing, we define (cx, cy, cz, c) as the same point as $(x,y,z,1)$. Those points of projective space are really lines through the origin in $\mathbb{R}^4$.
- Computer graphics use affine transformations (linear + shift). An affine transformation T is executed on $\mathbb{P}^3$ by a 4 by 4 matrix

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

shows translation

- Rotation
  - A rotation in $\mathbb{R}^2$ and $\mathbb{R}^3$ is achieved by an orthogonal matrix Q. The determinant is 1.

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

  - The above matrix rotates around the origin. How do we rotate around an arbitrary point like $(4,5)$. We translate $(4,5)$ to $(0,0)$, rotate by θ, then translate $(0,0)$ back to $(4,5)$

$$vT_-RT_+ = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & -5 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 5 & 1 \end{bmatrix}$$

  - The center of rotation $(4,5,1)$ moves to $(0,0,1)$. Rotation doesn't change it. Then it moves back to $(4,5,1)$
  - In $\mathbb{R}^3$, we get

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

leaves z-axis alone

leaves origin alone

  - The rotation matrix around a unit vector $a = (a_1, a_2, a_3)$

$$Q = (\cos\theta) I + (1-\cos\theta) \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ a_1a_2 & a_2^2 & a_2a_3 \\ a_1a_3 & a_2a_3 & a_3^2 \end{bmatrix} - \sin\theta \begin{bmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} & Q & & 0 \\ & & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Projection
  - A plane through the origin in a vector space, other planes are affine spaces, aka "flats". We want to project three-dimensional vectors onto planes. Start with a plane through the origin, whose normal vector is n. The vectors in the plane satisfy $n^T v = 0$. The usual projection onto the plane is $I - nn^T$
  - In homogeneous coordinates, this becomes 4 by 4

    projection onto the plane $n^T v = 0$ $\quad P = \begin{bmatrix} & & & 0 \\ & I - nn^T & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

  - Now, project onto a flat $n^T(v - v_0) = 0$, with v_0 as a point on the plane. First we translate v_0 to the origin by T_-, project it along the n direction, and translate back along the row vector v_0:

    Projection onto a flat $\quad T_- P T_+ = \begin{bmatrix} I & 0 \\ -v_0 & 1 \end{bmatrix} \begin{bmatrix} I - nn^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ v_0 & 1 \end{bmatrix}$
- Reflection
  - A reflection simply moves a point in the direction of a projection, but twice as far
  - Instead of $I - nn^T$, we have $I - 2nn^T$

    Reflection across a flat $\quad T_- R T_+ = \begin{bmatrix} I & 0 \\ -v_0 & 1 \end{bmatrix} \begin{bmatrix} I - 2nn^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ v_0 & 1 \end{bmatrix}$
- The projection matrix P gives a parallel projection. All points move parallel to n until they reach the plane. Another choice in computer graphics is a perspective projection, in which objects get smaller with distance.

Chapter 10: Complex Vectors and Matrices

Chapter 10.1: Complex Numbers

- $i = \sqrt{-1}$
- $Re(a+bi) = a$, $Im(a+bi) = b$
- $\bar{z}_1 \bar{z}_2 = \overline{z_1 z_2}$
- $\bar{z}_1 + \bar{z}_2 = \overline{z_1 + z_2}$
- If $Ax = \lambda x$ and A is real, then $A\bar{x} = \bar{\lambda}\bar{x}$
- The number $z = a + bi$ is also $z = r\cos\theta + ir\sin\theta = re^{i\theta}$
- Powers and multiplication is easy with complex numbers
- $z^n = r^n(\cos n\theta + i\sin n\theta)$
- Set $w = e^{2\pi i/n}$. The nth powers of $1, w, w^2, \ldots, w^{n-1}$ all equal 1