Last time:

- Markov matrices
- applications of matrices

Today:

- Multivariable chain rule
- Inverses of matrices

Problem 1: Chain Rule I

Define the functions $f : \mathbf{R}^2 \to \mathbf{R}^2$, $g : \mathbf{R}^3 \to \mathbf{R}$, and $h : \mathbf{R} \to \mathbf{R}^2$ by

$$f(y,z) = \begin{bmatrix} yz + e^y \\ \ln(1 + y^2 z^2) \end{bmatrix}, \quad g(r,s,t) = rs + t, \quad h(v) = \begin{bmatrix} ve^{-v} \\ v^2 \end{bmatrix}.$$

(a) Compute all three derivative matrices $(Df)(y,z)$, $(Dg)(r,s,t)$, and $(Dh)(v)$; make sure your matrix has the correct number of rows and columns in each case. Also compute $(Df)(1,1)$.

(b) Compute $(D(h \circ g))(1, -1, 1)$ in two ways: the Chain Rule, and by explicit computation of $(h \circ g)(r, s, t)$.

$$g(r, s, t) = rs + t, \quad h(v) = \begin{bmatrix} ve^{-v} \\ v^2 \end{bmatrix}.$$

(c) Compute the y-partial derivative of $g(y, f(y,z)) = g(y, f_1(y,z), f_2(y,z))$ in two ways: (i) work out $g(y, f(y,z))$ in terms of y and z, and (ii) use that $g(y, f(y,z)) = (g \circ q)(y,z)$ for $q(y,z) = (y, f(y,z)) = (y, yz + e^y, \ln(1+y^2z^2)) \in \mathbf{R}^3$ and compute $D(g \circ q)$ by the Chain Rule (in which the desired partial derivative is a specific matrix entry).

The first method is certainly easier in this case, so the point is just to see how the Chain Rule organizes the work very differently (its real power is for more complicated situations than this).

$$f(y,z) = \begin{bmatrix} yz + e^y \\ \ln(1+y^2z^2) \end{bmatrix}, \quad g(r,s,t) = rs + t, \quad h(v) = \begin{bmatrix} ve^{-v} \\ v^2 \end{bmatrix}$$

Problem 2: Chain Rule II

For a function $F(x, y)$, suppose $x = G(v, w)$ and $y = H(v, w)$ as expressed as functions of v and w, and that $v = k(r, s)$ and $w = \ell(r, s)$ are expressed as functions of r and s. Then $F(x, y)$ may be regarded as a function of r and s alone via such repeated substitutions. Explicitly:

$$F(x, y) = F(G(v, w), H(v, w)) = F(G(k(r, s), \ell(r, s)), H(k(r, s), \ell(r, s))) = \left(F \circ \begin{bmatrix} G \\ H \end{bmatrix} \circ \begin{bmatrix} k \\ \ell \end{bmatrix}\right)(r, s).$$

This comes up *all the time*: chains of dependencies of collections of variable on other collections of variables and so on.

Find an expression for $\dfrac{\partial F}{\partial r}$ in terms of partial derivatives of the functions: F (with respect to x and y), G and H (with respect to v and w), and k and ℓ (with respect to r and s).

Review: Matrix inverses

Problem 3: Computations with inverses

(a) Which of the following matrices are invertible? For each of the invertible ones, write down the inverse and check that it works by multiplying in both orders to confirm that you get I_2 each time:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & -4 \\ 5 & -3 \end{bmatrix}.$$

(b) For the matrices

$$A = \begin{bmatrix} 4 & 1 & -1 \\ 0 & 5 & 1 \\ 0 & 0 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 5 \\ 3 & -2 & 2 \end{bmatrix}$$

check that the matrices

$$A' = \begin{bmatrix} 1/4 & -1/20 & 1/20 \\ 0 & 1/5 & -1/30 \\ 0 & 0 & 1/6 \end{bmatrix}, \quad B' = \begin{bmatrix} 4 & -2 & 1 \\ 9 & -4 & 1 \\ 3 & -1 & 0 \end{bmatrix}$$

are respective inverses by multiplying to see that $A'A$ and AA' both equal I_3 and likewise for $B'B$ and BB'.

Problem 4: Using inversion

(a) Consider the system of equations

$$4x + y - z = 7, \;\; 5y + z = -3, \;\; 6z = 2.$$

Explain why this is the same as $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$, and A as in Problem 3(b). Then use your knowledge of A^{-1} from there to "solve for $\mathbf{x}$" (hint: multiply both sides of the vector formulation by A^{-1}), and check that the solution you obtained really works. Does the method work if the constants on the right side of these equations change?

(b) Consider the system of equations

$$3x + 2y = 7, \;\; 2x + y = -3, \;\; x + y = 2.$$

Explain why this is the same as $M\mathbf{x} = \mathbf{b}$ for $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$, and $M = \begin{bmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}$.

Verify that $M'M = \mathrm{I}_2$ for $M' = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -3 & 0 \end{bmatrix}$, and (via multiplication on the left by M') that if $M\mathbf{x} = \mathbf{b}$ then $\mathbf{x} = M'\mathbf{b}$. But compute $M'\mathbf{b}$ explicitly and check that it does *not* actually satisfy the given system of equations. What went wrong?

Problem 1: Chain Rule I

Define the functions $f : \mathbf{R}^2 \to \mathbf{R}^2$, $g : \mathbf{R}^3 \to \mathbf{R}$, and $h : \mathbf{R} \to \mathbf{R}^2$ by

$$f(y,z) = \begin{bmatrix} yz + e^y \\ \ln(1+y^2z^2) \end{bmatrix}, \quad g(r,s,t) = rs + t, \quad h(v) = \begin{bmatrix} ve^{-v} \\ v^2 \end{bmatrix}.$$

(a) Compute all three derivative matrices $(Df)(y,z)$, $(Dg)(r,s,t)$, and $(Dh)(v)$; make sure your matrix has the correct number of rows and columns in each case. Also compute $(Df)(1,1)$.

(b) Compute $(D(h \circ g))(1,-1,1)$ in two ways: the Chain Rule, and by explicit computation of $(h \circ g)(r,s,t)$.

(c) Compute the y-partial derivative of $g(y, f(y,z)) = g(y, f_1(y,z), f_2(y,z))$ in two ways: (i) work out $g(y, f(y,z))$ in terms of y and z, and (ii) use that $g(y, f(y,z)) = (g \circ q)(y,z)$ for $q(y,z) = (y, f(y,z)) = (y, yz + e^y, \ln(1+y^2z^2)) \in \mathbf{R}^3$ and compute $D(g \circ q)$ by the Chain Rule (in which the desired partial derivative is a specific matrix entry).

The first method is certainly easier in this case, so the point is just to see how the Chain Rule organizes the work very differently (its real power is for more complicated situations than this).

Problem 2: Chain Rule II

For a function $F(x,y)$, suppose $x = G(v,w)$ and $y = H(v,w)$ as expressed as functions of v and w, and that $v = k(r,s)$ and $w = \ell(r,s)$ are expressed as functions of r and s. Then $F(x,y)$ may be regarded as a function of r and s alone via such repeated substitutions. Explicitly:

$$F(x,y) = F(G(v,w), H(v,w)) = F(G(k(r,s), \ell(r,s)), H(k(r,s), \ell(r,s))) = \left(F \circ \begin{bmatrix} G \\ H \end{bmatrix} \circ \begin{bmatrix} k \\ \ell \end{bmatrix} \right)(r,s).$$

This comes up *all the time*: chains of dependencies of collections of variable on other collections of variables and so on.

Find an expression for $\dfrac{\partial F}{\partial r}$ in terms of partial derivatives of the functions: F (with respect to x and y), G and H (with respect to v and w), and k and ℓ (with respect to r and s).

Problem 3: Computations with inverses

(a) Which of the following matrices are invertible? For each of the invertible ones, write down the inverse and check that it works by multiplying in both orders to confirm that you get I_2 each time:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & -4 \\ 5 & -3 \end{bmatrix}.$$

(b) For the matrices

$$A = \begin{bmatrix} 4 & 1 & -1 \\ 0 & 5 & 1 \\ 0 & 0 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 5 \\ 3 & -2 & 2 \end{bmatrix}$$

check that the matrices

$$A' = \begin{bmatrix} 1/4 & -1/20 & 1/20 \\ 0 & 1/5 & -1/30 \\ 0 & 0 & 1/6 \end{bmatrix}, \quad B' = \begin{bmatrix} 4 & -2 & 1 \\ 9 & -4 & 1 \\ 3 & -1 & 0 \end{bmatrix}$$

are respective inverses by multiplying to see that $A'A$ and AA' both equal I_3 and likewise for $B'B$ and BB'.

Problem 4: Using inversion

(a) Consider the system of equations

$$4x + y - z = 7,\ \ 5y + z = -3,\ \ 6z = 2.$$

Explain why this is the same as $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$, and A as in Problem 3(b). Then use your knowledge of A^{-1} from there to "solve for $\mathbf{x}$" (hint: multiply both sides of the vector formulation by A^{-1}), and check that the solution you obtained really works. Does the method work if the constants on the right side of these equations change?

(b) Consider the system of equations

$$3x + 2y = 7,\ \ 2x + y = -3,\ \ x + y = 2.$$

Explain why this is the same as $M\mathbf{x} = \mathbf{b}$ for $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$, and $M = \begin{bmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}$.

Verify that $M'M = \mathrm{I}_2$ for $M' = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -3 & 0 \end{bmatrix}$, and (via multiplication on the left by M') that if $M\mathbf{x} = \mathbf{b}$ then $\mathbf{x} = M'\mathbf{b}$. But compute $M'\mathbf{b}$ explicitly and check that it does *not* actually satisfy the given system of equations. What went wrong?

