

1. Suppose that you are given \mathbf{A} , \mathbf{B} , and \vec{x} . Explain how to evaluate the expression $(\mathbf{A} + \mathbf{B}^{-T})^{-1}\vec{x}$ without computing the inverse of any matrices. **Solution:** Work backwards. Let \vec{y} denote the final vector

$$\vec{y} = (\mathbf{A} + \mathbf{B}^{-1})^{-1}\vec{x} \Leftrightarrow (\mathbf{A} + \mathbf{B}^{-1})\vec{y} = \vec{x}.$$

At this point if we knew \mathbf{B}^{-1} we could add it to \mathbf{A} and then compute \vec{y} by solving the linear system. But we still don't know \mathbf{B}^{-1} . Multiply the whole system by \mathbf{B} to get

$$(\mathbf{BA} + \mathbf{I})\vec{y} = \mathbf{B}\vec{x}.$$

This expression only involves matrices that we know, so we can find \vec{y} using the following sequence of steps:

- Compute $\vec{z} = \mathbf{B}\vec{x}$.
 - Compute the coefficient matrix $\mathbf{C} = \mathbf{BA} + \mathbf{I}$.
 - Solve the linear system $\mathbf{C}\vec{y} = \vec{z}$.
2. Without using the Wronskian, prove that the following three functions are linearly independent

$$p_1(x) = 1, p_2(x) = 1 - x, p_3(x) = 1 - x + x^2.$$

Solution: A linear combination of these three ‘vectors’ takes the form

$$c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x) = c_1 + c_2 + c_3 + (-c_2 - c_3)x + (c_3)x^2.$$

This polynomial is equal to the zero polynomial only when all of the coefficients are zero, i.e. when the coefficients satisfy

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \vec{0}.$$

The matrix is already in row echelon form and has full column rank. The only solution is the trivial solution, which implies that the three polynomials form a basis for their span, which is the space of polynomials of degree ≤ 2 .

The foregoing is an acceptable solution; from here on we just provide extra info. When we say ‘This polynomial is equal to the zero polynomial only when all of the coefficients are zero’ we are implicitly using the monomial basis 1 , x , and x^2 , and we are implicitly assuming that these functions/vectors are linearly independent. One way of showing that a set of functions is linearly independent is to use the Wronskian. Another way is to find a set of n points x_n , evaluate the n functions at those points to produce vectors, and then show that the vectors are linearly independent. The idea is that if it's not possible to find a nontrivial linear combination of the functions that is zero *at these points* then clearly it's not possible to find a nontrivial linear combination of the functions that is zero *everywhere*. A drawback to this latter approach is that you could find a set of points where the vectors are dependent, but the functions are still independent.

For example, suppose that we want to show that 1 , x , and x^2 are linearly independent. We choose three points $x_1 = -1$, $x_2 = 0$, and $x_3 = 1$. We evaluate the functions at these points to produce vectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

You can show that these vectors are linearly independent, which means that the functions that produced the vectors are also independent. In fact, for these three functions you can show that as long as the three points x_1 , x_2 , and x_3 are distinct, then the vectors produced in this way are linearly independent.

For other functions it's not so simple. For example, the functions $\cos(t)$, $\sin(t)$, and $\sin(2t)$ are linearly independent, but if you choose the points $0, \pi/2, \pi$ then the vectors will be linearly dependent. So you sometimes have to choose your points carefully.

3. Consider the vector spaces of polynomials V of degree at most 2 and W of degree at most 3.

- (a) Find the matrix representation of the following linear function from V to W

$$L[p] = 2x \frac{dp}{dx}$$

with the monomial basis for both spaces. **Solution:** Recall that the entries of the matrix are such that

$$L[\vec{v}_j] = \sum_i a_{ij} \vec{w}_i.$$

Applying L to the monomial basis for V yields

$$L[1] = 2x \times 0 = 0 \times 1 + 0 \times x + 0 \times x^2 + 0 \times x^3$$

$$L[x] = 2x \times 1 = 0 \times 1 + 2 \times x + 0 \times x^2 + 0 \times x^3$$

$$L[x^2] = 2x \times (2x) = 0 \times 1 + 0 \times x + 4 \times x^2 + 0 \times x^3$$

The matrix is therefore

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

- (b) Find the matrix representation of the same linear function with respect to the Chebyshev basis for both spaces: $1, x, 2x^2 - 1$, and for W also $4x^3 - 3x$. **Solution:** As above, we have

$$L[1] = 2x \times 0 = a_{1,1} \times 1 + a_{2,1} \times x + a_{3,1} \times (2x^2 - 1) + a_{4,1} \times (4x^3 - 3x)$$

$$L[x] = 2x \times 1 = a_{1,2} \times 1 + a_{2,2} \times x + a_{3,2} \times (2x^2 - 1) + a_{4,2} \times (4x^3 - 3x)$$

$$L[2x^2 - 1] = 2x \times (4x) = a_{1,3} \times 1 + a_{2,3} \times x + a_{3,3} \times (2x^2 - 1) + a_{4,3} \times (4x^3 - 3x)$$

This time we have to do a little extra work, but in the end we find

$$\begin{bmatrix} 0 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

4. Consider the function $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that maps the initial condition to the final condition at time $t = \tau$ for a linear, non-autonomous, homogeneous system of differential equations:

$$\frac{d\vec{x}}{dt} = \mathbf{A}(t)\vec{x}, \quad \vec{x}(0) = \vec{x}_0.$$

- (a) Prove that this function is linear. **Solution:** Check the two conditions for linearity.

- Suppose that we have two initial conditions \vec{x}_0 and \vec{y}_0 that lead to two solutions $\vec{x}(t)$ and $\vec{y}(t)$ that satisfy

$$\frac{d\vec{x}}{dt} = \mathbf{A}(t)\vec{x}, \quad \vec{x}(0) = \vec{x}_0,$$

$$\frac{d\vec{y}}{dt} = \mathbf{A}(t)\vec{y}, \quad \vec{y}(0) = \vec{y}_0.$$

Now suppose we add the initial conditions and get $\vec{z}(t)$ as the solution to

$$\frac{d\vec{z}}{dt} = \mathbf{A}(t)\vec{z}, \quad \vec{z}(0) = \vec{x}_0 + \vec{y}_0.$$

Do we have $\vec{z}(t) = \vec{x}(t) + \vec{y}(t)$? Yes, we do. Adding the ODEs for \vec{x} and \vec{y} produces the ODE for \vec{z} .

- Suppose that we multiply the initial condition \vec{x}_0 by a constant c to get a new initial condition $\vec{y}_0 = c\vec{x}_0$, then define \vec{y} to be the solution of

$$\frac{d\vec{y}}{dt} = \mathbf{A}(t)\vec{y}, \quad \vec{y}(0) = \vec{y}_0.$$

Is $\vec{y}(t) = c\vec{x}(t)$? Yes, it is. Multiplying the ODE for \vec{x} by c produces the ODE for \vec{y} .

- (b) Infer from (a) that $\vec{x}(t) = \Pi(t)\vec{x}_0$ (no work required). Write down a system of ODEs that $\Pi(\tau)$ must solve (explain why).

Solution: In the general case we are looking for a matrix \mathbf{A} whose entries are the solution of $L[\vec{v}_j] = \sum_i a_{ij}\vec{w}_i$. In this specific case we re-named the matrix from \mathbf{A} to $\mathbf{\Pi}$ because we were using the letter \mathbf{A} for something else. Let the entries of $\mathbf{\Pi}$ be denoted π_{ij} , and note that they depend on which final time t you're interested in.

In this specific case $V = W = \mathbb{R}^n$ and we are using the standard basis, so the entries of $\mathbf{\Pi}$ solve

$$L[\vec{e}_j] = \sum_i \pi_{ij}\vec{e}_i = \begin{pmatrix} \pi_{1j} \\ \vdots \\ \pi_{nj} \end{pmatrix}.$$

We don't have a nice explicit expression for these entries of the matrix $\mathbf{\Pi}$. But we do know that the vector on the right of the equation above is obtained by solving the ODE with initial condition \vec{e}_j , i.e. we know

$$\frac{d}{dt} \begin{pmatrix} \pi_{1j} \\ \vdots \\ \pi_{nj} \end{pmatrix} = \mathbf{A}(t) \begin{pmatrix} \pi_{1j} \\ \vdots \\ \pi_{nj} \end{pmatrix}, \quad \begin{pmatrix} \pi_{1j}(t=0) \\ \vdots \\ \pi_{nj}(t=0) \end{pmatrix} = \vec{e}_j.$$

If we stack these vector equations left to right we get the following matrix equation

$$\left[\frac{d}{dt} \begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{n1} \end{pmatrix} \cdots \frac{d}{dt} \begin{pmatrix} \pi_{1n} \\ \vdots \\ \pi_{nn} \end{pmatrix} \right] = \left[\mathbf{A}(t) \begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{n1} \end{pmatrix} \cdots \mathbf{A}(t) \begin{pmatrix} \pi_{1n} \\ \vdots \\ \pi_{nn} \end{pmatrix} \right]$$

or in more compact notation

$$\frac{d\mathbf{\Pi}}{dt} = \mathbf{A}(t)\mathbf{\Pi}.$$

We also have an initial condition for each system of vector ODEs, and by stacking them left to right we get an initial condition for the matrix system of ODEs:

$$\left[\begin{pmatrix} \pi_{11}(t=0) \\ \vdots \\ \pi_{n1}(t=0) \end{pmatrix} \cdots \begin{pmatrix} \pi_{1n}(t=0) \\ \vdots \\ \pi_{nn}(t=0) \end{pmatrix} \right] = [\vec{e}_1 \cdots \vec{e}_n]$$

or in more compact notation

$$\mathbf{\Pi}(t=0) = \mathbf{I}.$$

5. Let $\vec{0} \neq \vec{v} \in \mathbb{R}^m$ and $\vec{0} \neq \vec{w} \in \mathbb{R}^n$. Prove that the matrix $\vec{v}\vec{w}^T$ has rank one.

Solution There are lots of ways to think about rank. For this particular matrix, we will symbolically compute the Row Echelon Form and then count the number of pivots.

Permutation matrices are invertible, so we can multiply from the left and/or the right by a perturbation matrix without changing the rank. This implies that we can assume that $v_1 \neq 0$ and $w_1 \neq 0$ without loss of generality. Now consider

$$\vec{v}\vec{w}^T = \begin{bmatrix} v_1w_1 & v_1w_2 & \cdots & v_1w_n \\ v_2w_1 & v_2w_2 & \cdots & v_2w_n \\ \vdots & & & \vdots \\ v_mw_1 & v_mw_2 & \cdots & v_mw_n \end{bmatrix}.$$

Eliminate the (2,1) entry by multiplying row #1 by $-v_2/v_1$ and adding to row 2:

$$\vec{v}\vec{w}^T = \begin{bmatrix} v_1w_1 & v_1w_2 & \cdots & v_1w_n \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ v_mw_1 & v_mw_2 & \cdots & v_mw_n \end{bmatrix}.$$

Eliminating all the remaining entries below the diagonal in the first column leads to

$$\vec{v}\vec{w}^T = \begin{bmatrix} v_1w_1 & v_1w_2 & \cdots & v_1w_n \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The REF only has one pivot so the rank is one.

6. Find the rank of this $n \times n$ matrix

$$\mathbf{I} - \frac{1}{n}\mathbf{1}$$

where $\mathbf{1}$ is a matrix whose every entry is 1.

Solution: There are lots of ways to think about rank. For this matrix we will find the dimension of the null space, then say that the rank equals the number of columns minus the dimension of the null space.

Vectors in the null space satisfy

$$\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\right)\vec{z} = \vec{0}$$

which can be re-arranged to

$$\vec{z} = \frac{1}{n}\mathbf{1}\vec{z}.$$

Now consider what happens when you multiply the one matrix by \vec{z} : you get a vector whose entries are all equal to the sum of the entries of \vec{z} . Denote the *average* of the entries of \vec{z} by \bar{z} . The above equation becomes

$$\vec{z} = \bar{z}\vec{1}.$$

All vectors in the null space must therefore be parallel to $\vec{1}$, so the null space is one-dimensional. The rank of the original matrix is therefore $n - 1$.

7. (Sherman-Morrison) Let \mathbf{A} be an $n \times n$ invertible matrix and let $\vec{u}, \vec{v} \in \mathbb{R}^n$. Without even *thinking* about using a determinant, prove

(a) that $\mathbf{A} + \vec{u}\vec{v}^T$ is invertible iff $1 + \vec{v}^T\mathbf{A}^{-1}\vec{u} \neq 0$, and

(b) in that case

$$(\mathbf{A} + \vec{u}\vec{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{1}{1 + \vec{v}^T\mathbf{A}^{-1}\vec{u}}\mathbf{A}^{-1}\vec{u}\vec{v}^T\mathbf{A}^{-1}.$$

(Hint: Use the expression given in (b) to prove one direction in (a). To prove the other direction, consider the vector $\mathbf{A}^{-1}\vec{u}$.)

Solution: First assume $1 + \vec{v}^T\mathbf{A}^{-1}\vec{u} \neq 0$ and just verify the expression given for the inverse in (b). This will prove one direction of (a):

$$\begin{aligned} (\mathbf{A} + \vec{u}\vec{v}^T) \left(\mathbf{A}^{-1} - \frac{1}{1 + \vec{v}^T\mathbf{A}^{-1}\vec{u}}\mathbf{A}^{-1}\vec{u}\vec{v}^T\mathbf{A}^{-1} \right) = \\ \mathbf{I} - \frac{1}{1 + \vec{v}^T\mathbf{A}^{-1}\vec{u}}\vec{u}\vec{v}^T\mathbf{A}^{-1} + \vec{u}\vec{v}^T\mathbf{A}^{-1} - \frac{1}{1 + \vec{v}^T\mathbf{A}^{-1}\vec{u}}\vec{u}\vec{v}^T\mathbf{A}^{-1}\vec{u}\vec{v}^T\mathbf{A}^{-1} \end{aligned}$$

In the last expression, note that $\vec{v}^T \mathbf{A}^{-1} \vec{u}$ is a scalar and can be factored out:

$$\vec{u} \vec{v}^T \mathbf{A}^{-1} \vec{u} \vec{v}^T \mathbf{A}^{-1} = \vec{u} (\vec{v}^T \mathbf{A}^{-1} \vec{u}) \vec{v}^T \mathbf{A}^{-1} = (\vec{v}^T \mathbf{A}^{-1} \vec{u}) \vec{u} \vec{v}^T \mathbf{A}^{-1}$$

Plugging this back in, we have

$$\begin{aligned} (\mathbf{A} + \vec{u} \vec{v}^T) \left(\mathbf{A}^{-1} - \frac{1}{1 + \vec{v}^T \mathbf{A}^{-1} \vec{u}} \mathbf{A}^{-1} \vec{u} \vec{v}^T \mathbf{A}^{-1} \right) = \\ \mathbf{I} + \left[-\frac{1}{1 + \vec{v}^T \mathbf{A}^{-1} \vec{u}} + 1 - \frac{\vec{v}^T \mathbf{A}^{-1} \vec{u}}{1 + \vec{v}^T \mathbf{A}^{-1} \vec{u}} \right] \vec{u} \vec{v}^T \mathbf{A}^{-1} \end{aligned}$$

The term in square brackets is zero. This proves half of (a) and (b). If you prefer you can reverse the order and verify that way using similar manipulations; there is no need to do both because an inverse matrix is an inverse from the left and from the right.

To prove the remaining part of (a), we need to assume that \mathbf{A} is invertible and that $\mathbf{A} + \vec{u} \vec{v}^T$ is also invertible and show that this implies that $1 + \vec{v}^T \mathbf{A}^{-1} \vec{u} \neq 0$. Per the hint, consider what happens if we apply $\mathbf{A} + \vec{u} \vec{v}^T$ to the vector $\mathbf{A}^{-1} \vec{u}$:

$$(\mathbf{A} + \vec{u} \vec{v}^T) \mathbf{A}^{-1} \vec{u} = \vec{u} + \vec{u} \vec{v}^T \mathbf{A}^{-1} \vec{u} = \vec{u} + \vec{u} (\vec{v}^T \mathbf{A}^{-1} \vec{u}) = (1 + \vec{v}^T \mathbf{A}^{-1} \vec{u}) \vec{u}.$$

If $1 + \vec{v}^T \mathbf{A}^{-1} \vec{u} = 0$ then the vector $\mathbf{A}^{-1} \vec{u}$ is in the null space of $\mathbf{A} + \vec{u} \vec{v}^T$, which would contradict the assumption that the $\mathbf{A} + \vec{u} \vec{v}^T$ is invertible (because the only vector in the null space of an invertible matrix is $\vec{0}$).

Comment: This is useful in a context where you can easily solve linear systems with coefficient matrix \mathbf{A} , e.g. because you have an LU factorization of \mathbf{A} . The Sherman-Morrison identity allows you to find the solution of linear systems whose coefficient matrix is a rank-one perturbation of \mathbf{A} without having to row reduce the new coefficient matrix. Using the formula in part (b) only requires solving linear systems with coefficient matrix \mathbf{A} (and also taking some dot products, which is easy).

8. Prove that if $\vec{v}_1, \dots, \vec{v}_r$ are a basis for the corange of \mathbf{A} , then $\mathbf{A} \vec{v}_1, \dots, \mathbf{A} \vec{v}_r$ are a basis for the range of \mathbf{A} .

Solution: The dimension of the range equals that of the corange, and the vectors $\mathbf{A} \vec{v}_1, \dots, \mathbf{A} \vec{v}_r$ are all in the range by construction. All we need to prove is that they are linearly independent. So we ask whether there are coefficients c_i , not all zero, such that

$$c_1 \mathbf{A} \vec{v}_1 + \dots + c_r \mathbf{A} \vec{v}_r = \vec{0}.$$

Simplifying,

$$\mathbf{A}(c_1 \vec{v}_1 + \dots + c_r \vec{v}_r) = \vec{0}.$$

The above equation is true when either $c_1 \vec{v}_1 + \dots + c_r \vec{v}_r = \vec{0}$, or $c_1 \vec{v}_1 + \dots + c_r \vec{v}_r$ is in the null space of \mathbf{A} . The first condition is not possible because the vectors are independent, and the second is not possible because the vectors are all in the corange. (The corange and null space intersect only at $\vec{0}$.) So the only solution is for all the coefficients c_i to be zero, which means the vectors are independent, which means they are a basis.