Inverse matrix, rank

Today we will be discussing inverse matrices and representation of linear maps with respect to different bases. To begin, let $A$ be an $n \times n$ matrix.

**Definition 6.1.** Write $B = A^{-1}$ if $AB = I_n$.

**Exercise 6.2.** If $AB = I_n$, then $BA = I_n$. We call such a $B$ an inverse for $A$.

Now suppose that $A = (a_1, \ldots, a_n)$ is a $k \times n$ matrix over the field $\mathbb{F}$, that is $A \in \mathbb{F}^{k \times n}$. If $b \in \mathbb{F}^k$, then $\exists x \in \mathbb{F}^n$ such that $Ax = b \iff b \in \text{Span}(a_1, \ldots, a_n) = \text{column space of } A$. If $R$ is an $n \times k$ matrix, then $AR = A[r_1, \ldots, r_k] = [Ar_1, \ldots, Ar_k]$. This leads us to the following definition.

**Definition 6.3.** $R$ is a right inverse of $A$ if $AR = I_k$.

We are naturally led to wonder when a right inverse should exist. Letting $\{e_1, \ldots, e_k\}$ be the standard basis for $\mathbb{F}^k$, we see that $R$ exists $\iff \forall i \exists r_i$ such that $Ar_i = e_i$. This condition is equivalent to $e_1, \ldots, e_k \in \text{column space of } A \leq \mathbb{F}^k$. This means that the column space of $A$ is equal to $\mathbb{F}^k$; in other words, $A$ has rank $k$. So, a $k \times n$ matrix $A$ has a right inverse $\iff \text{rk } A = k$. We restate our findings as a theorem.

**Theorem 6.4.** The following are equivalent for a $k \times n$ matrix $A$:

1. $A$ has a right inverse.
2. $\text{rk } A = k$.
3. $A$ has full row rank.
4. $A$ has linearly independent rows.
Remember that if $A = (a_{ij})$, then the transpose of $A$ is the matrix given by $A^T = (a_{ji})$.

**Exercise 6.5.** We have the formula $(AB)^T = B^T A^T$.

Using the exercise, we see that $AR = I \iff R^T A^T = I$. Thus, $A$ has a right inverse iff $A^T$ has a left inverse. We therefore have a new theorem.

**Theorem 6.6.** The following are equivalent for a $k \times n$ matrix $A$:

1. $A$ has a left inverse.
2. $\text{rk} A = n$.
3. $A$ has full column rank.
4. $A$ has linearly independent columns.

The set of $n \times n$ matrices over a field $\mathbb{F}$ is very important and has its own notation, $M_n(\mathbb{F})$. In this case, the previous two theorems coincide.

**Corollary 6.7.** The following are equivalent for $A \in M_n(\mathbb{F})$:

1. $\text{rk} A = n$.
2. $A$ has a right inverse.
3. $A$ has a left inverse.
4. $A$ has an inverse.
5. $\det A \neq 0$ ($A$ is nonsingular).

**Exercise 6.8.** Let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{F}$ be all distinct. Now let $\gamma_{i,j} = 1/(\alpha_i - \beta_j)$. The matrix $H = (\gamma_{i,j})$ has full rank. Matrices such as $H$ are called Cauchy-Hilbert matrices.

The set of matrices considered in the previous corollary pervades the study of linear algebra, so we give it a name.

**Definition 6.9.** The set of nonsingular $n \times n$ matrices over $\mathbb{F}$ is called the **General Linear Group**, and is denoted by $GL(n, \mathbb{F})$.

To justify the name, notice that $GL(n, \mathbb{F})$ is a group under matrix multiplication. Only two axioms require effort to check. First, see that $(A^{-1})^{-1} = A$, so $GL(n, \mathbb{F})$ is closed under taking inverses. Second, see that if $A, B \in GL(n, \mathbb{F})$, then $(AB)^{-1} = B^{-1}A^{-1}$. Therefore $GL(n, \mathbb{F})$ is closed under the group operation. Associativity and the existence of an identity element are clear, so we see that the general linear group is indeed a group.
Now we can use our understanding of matrix inversion to learn about changes of basis. Let \( \varphi : V^n \to W^k \) be a linear map, and suppose we have two bases for each vector space: \( e_i, e'_i; f_j, f'_j \).

Now consider the basis change transformations

\[
\begin{align*}
\sigma : V &\to V, & \sigma(e_i) &= e'_i \\
\tau : W &\to W, & \tau(f_i) &= f'_i
\end{align*}
\]

Define \( S := [\sigma]_e^e \) and \( T := [\tau]_f^f \). Similarly, let \( S' := [\sigma]_{e'}^{e'} \) and \( T' := [\tau]_{f'}^{f'} \). Notice that all four of these matrices are nonsingular because their columns are vector space bases. Now define the matrices \( A = [\varphi]_e^f \) and \( A' = [\varphi]_{e'}^{f'} \). Note that if \( x \) is a column vector in \( V \), then \( [\varphi x]_f = [\varphi]_e^f [x]_e \).

Our first goal is to compare \( u = [x]_e \) with \( u' = [x]_{e'} \). Let’s write \( u = u_1 e_1 + \ldots + u_n e_n \). Now consider the following simple and surprising calculation:

\[
u' = \sigma x = \sigma(\sum u_i e_i) = \sum u_i \sigma(e_i) = \sum u_i e'_i.
\]

This tells us that \( u = [x]_e = [\sigma x]_{e'} = [\sigma]_{e}^{e'} [x]_e = S'[x]_e = S' u' \).

So, \( u = S' u' \) and \( u' = (S')^{-1} u \), accomplishing our first goal.

Now we can turn to our second goal, which is to compare \( A \) with \( A' \). Define \( v = A u \) and \( v' = A' u' \). Now we can see that

\[
T' v' = v = A u = AS' u' .
\]

In other words,

\[
(T')^{-1} AS' u' = v' = A' u' .
\]

Therefore, we have the formula

\[
A' = (T')^{-1} AS' .
\]

We can actually clean this formula up a bit by considering the case where \( A = S \) and \( A' = S' \). In this case, \( \tau = \sigma \), so what above were \( T \) and \( T' \) are now \( S \) and \( S' \). So the formula now reads: \( S' = (S')^{-1} S S' \). Multiplying on the right by \( (S')^{-1} \) then on the left by \( S' \), we find that \( S' = S \). We could do the same thing with \( T \) to find that \( T' = T \), so our nicer formula has the form:

\[
A' = T^{-1} A S.
\]

**Exercise 6.10.** If \( A \) is nonsingular, then \( \text{rk}(AB) = \text{rk} B \) and \( \text{rk}(CA) = \text{rk} C \).

**Exercise 6.11.** \( \text{rk}(AB) \leq \max\{\text{rk} A, \text{rk} B\} \).

**Exercise 6.12.** \( \text{rk}(A + B) \leq \text{rk} A + \text{rk} B \).
2 Similarity of matrices, characteristic polynomial

Let $A$ be an $n \times n$ matrix representing a linear map $V \to V$. Such a linear map is called a linear transformation. A change of basis matrix $S \in GL(n, \mathbb{F})$ represents a linear transformation. If $A$ and $A'$ represent the same linear transformation with respect to the two bases between which $S$ changes, then we have $A' = S^{-1}AS$. This important concept leads us to a definition.

Definition 6.13. If $A, B \in M_n(\mathbb{F})$, then they are similar (or conjugate) if $\exists S \in GL(n, \mathbb{F})$ such that $B = S^{-1}AS$. This is denoted by $A \sim B$.

Theorem 6.14. Let $V$ be a vector space and $\varphi$ a linear transformation. Then for any two bases $(e_i, e'_i)$, $[\varphi]_e \sim [\varphi]_{e'}$.

Exercise 6.15. Similarity of matrices is an equivalence relation.

Recall the determinant function $\det : M_n(\mathbb{F}) \to \mathbb{F}$.

Exercise 6.16. $\det(AB) = \det A \det B$.

We have a neat formula for the determinant of an inverse matrix. Consider

$$AA^{-1} = I \Rightarrow \det(\det AA^{-1}) = \det I = 1.$$ 

Then, $\det(\det AA^{-1}) = \det A \det A^{-1} \Rightarrow \det A^{-1} = 1/\det A$.

Exercise 6.17. If $A \sim B$, then $\det A = \det B$.

Now recall that for an $n \times n$ matrix $A$, the trace of $A$ is given by the formula $\text{tr} A = \sum_{i=1}^{n} a_{ii}$.

Exercise 6.18. For $A \in \mathbb{F}^{k \times n}$ and $B \in \mathbb{F}^{n \times k}$, we have $\text{tr}(AB) = \text{tr}(BA)$.

Now let $A \in M_n(\mathbb{F})$ and $x$ be a variable in $\mathbb{F}$.

Definition 6.19. The characteristic matrix of $A$ is the matrix $xI - A$. The characteristic polynomial of $A$ is the polynomial $f_A(x) := \det(xI - A)$.

Example 6.20. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, so $\det A = -2$ and $\text{tr} A = 5$. Then the characteristic matrix of $A$ is $xI - A = \begin{pmatrix} x - 1 & -2 \\ -3 & x - 4 \end{pmatrix}$. Then the characteristic polynomial of $A$ is $f_A(x) = \begin{vmatrix} x - 1 & -2 \\ -3 & x - 4 \end{vmatrix} = (x - 1)(x - 4) - 6 = x^2 - 5x - 2 = x^2 - \text{tr} A + \det A$.

Exercise 6.21. The characteristic polynomial of $A$ is actually given by

$$f_A(x) = x^n - \text{tr} A x^{n-1} + \cdots + (-1)^n \det A.$$ 

Exercise 6.22. If $A \sim B$, then $f_A(x) = f_B(x)$. 


Since matrices which represent the same linear map with respect to different bases are similar, we can make the following definition.

**Definition 6.23.** Let $\varphi : V \to V$ be linear. The **characteristic polynomial** of $\varphi$ is given by $f_\varphi(x) := f_A(x)$, where $A = [\varphi]$ in some basis.

Finally, let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

**Exercise 6.24.** Calculate $f_A(A) = A^2 - (a + d)A + (ad - bc)I$ to find a curious result.