UNDERSTADING LINEAR MAPS WITH GAUSS-JORDAN ELIMINATION

Abstract. We provide a condensed proof of the fundamental theorem of linear algebra in a Gauss-Jordan elimination fashion.

1.1. We denote $\mathbb{R}^{m\times n}$ the collection of all matrices of shape $m \times n$.

1.2. Any matrix $A \in \mathbb{R}^{m\times n}$ is associated with a linear map that we can denote $f_A : \mathbb{R}^n \to \mathbb{R}^m$.

1.3. Column-wise elementary operations can be of either of three kinds:

(1) Permutation of columns.
(2) Replacing a column c_i with λc_i with $\lambda \neq 0$.
(3) Replacing a column c_i with $c_i + \mu c_j$ for any μ.

We denote $e(A)$ the transform of a matrix A by a given elementary operation e.

1.4. Hereinafter our discussion will be based entirely on column-wise operations, so we will drop "column-wise" and "column" from all our statements whenever possible.

1.5. Each elementary operation of A is equivalent to multiplying A by an appropriate matrix $E \in \mathbb{R}^{n\times n}$. Hence, we say that E is the matrix representing the elementary operation e whenever $e(A) = AE$ for all possible matrices A.

1.6. Note that elementary operations preserve the rank, i.e., $\operatorname{rank}(A) = \operatorname{rank}(e(A)) = \operatorname{rank}(AE)$. Can you see why?

1.7. In Gauss-Jordan elimination we always keep track of two matrices, that we can simply represent as a tuple (A, B) where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times n}$. We convene $(A, B) \sim (A', B')$ to mean that (A', B') results from (A, B) by applying the same sequence of elementary operations to both matrices.

1.8. Gauss-Jordan elimination is the method whereby the pair $(A, \mathrm{Id}_{n\times n})$ is transformed, by means of applying a sequence of elementary operations, into another pair (L, B) where L has a canonical form known as "reduced echelon form". This means that L satisfies the following requirements:

(1) L is a lower column echelon matrix.
(2) The leading entries of L are 1.
(3) The leading entries of L are the only non-zero entries in their row.

Following our notation: $(A, \mathrm{Id}_{n\times n}) \sim (L, B)$.

1.9. Given an input matrix A, Gauss-Jordan elimination leads to a unique matrix L in reduced echelon form. We can stress this fact by defining a Gauss-Jordan algorithm $\mathrm{GJ} : A \mapsto L(A)$ that accepts A as input and returns its unique reduced column echelon form.

1.10. Proposition. If $(A, \mathrm{Id}_{n\times n}) \sim (L, B)$ then $L = AB$.

Proof. Let $E = E_1 \cdots E_k$ the product of the matrices encoding the elementary operations that have been applied. By definition $L = AE$ and $B = \mathrm{Id}_{n\times n}E = E$, so it is clear that $B = E$, i.e., B keeps track of all the elementary operations that have been applied to A that give L. It follows that $L = AB$. □

1.11. Observe that $\mathrm{rank}(B) = n$. Can you figure out why?

1.12. Proposition. If A is a square, full-rank matrix, then $L(A) = \mathrm{Id}_{n\times n}$ and $B = A^{-1}$.

Proof. If $(A, \mathrm{Id}_{n\times n}) \sim (\mathrm{Id}_{n\times n}, B)$ for some matrix B, then we know by the previous discussion that $\mathrm{Id}_{n\times n} = AB$, so $B = A^{-1}$. □

1.13. For the next discussion, let $(A, \mathrm{Id}_{n\times n}) \sim (L, B)$, with L in reduced echelon form. Let $B = [\ B_1 \mid \ldots \mid B_n\ ]$ and $L = [\ L_1 \mid \ldots \mid L_r \mid 0 \mid \ldots \mid 0\ ]$ specified by their respective columns, where $L_1, \ldots, L_r$ are non-zero column vectors.

1.14. Theorem. $\mathrm{rank}(A) = r$, where r is the number of non-zero columns in L. Moreover, $L_1, \ldots, L_r$ form a basis of the column space of A, denoted $C(A)$. Therefore, $\dim C(A) = r$.

1.15. Observe that the last $n - r$ columns of L are zero. What does this mean?

1.16. Theorem. $B_1, \ldots, B_{n-r}$ form a basis of the null-space of A, denoted $N(A)$.

Proof. The following are known facts:

(1) $B_1, \ldots, B_n$ is a basis of $\mathbb{R}^n$; in particular, any subset is linearly independent.
(2) The vectors $L_i = AB_i$ for each $1 \leq i \leq r$ form a linearly independent set.
(3) The vectors $L_i = AB_i$ are zero for $r + 1 \leq i \leq n$. Consequently,

$$S = \mathrm{span}\{B_i \mid r + 1 \leq i \leq n\}\} \subset N(A).$$

Let's check that $N(A) \subset S$. By (1) we can write any $v \in N(A)$ as a linear combination $v = \sum_{i=1}^{n} \lambda_i B_i$. From the fact that $v \in N(A)$ and (3), it follows that $0 = Av = \sum_{i=1}^{r} \lambda_i AB_i = \sum_{i=1}^{r} \lambda_i L_i$. By (2) this cannot hold unless $\lambda_i = 0$ for $i = 1, \ldots, r$. Then $v \in S$. □

1.17. It follows that $\dim N(A) = n - r$.

1.18. Fundamental Theorem of Linear Algebra. For any matrix $A \in \mathbb{R}^{m\times n}$ the following identity holds:

$$\dim N(A) + \dim C(A) = n.$$