Five Proofs of the Subsequence Theorem
An Exposition By William Gasarch

0.1 Introduction

The following theorem was first proven by Erdös and Szekeres [2].

Theorem 0.1.1 *Let* $x_1, \ldots, x_{n^2+1}$ *be any sequence of distinct reals. Then there exists either an increasing or decreasing* $(n+1)$*-subsequence.*

We refer to this as *the subsequence theorem* throughout.

We present five (known) proofs of the subsequence theorem. We first tell how we were lead to these proofs. After Martin Kruskal died, Clyde Kruskal (his son) was looking through his papers. He found a manuscript, dated 1950, by Joseph Kruskal that discusses the subsequence theorem and some multidimensional versions of it. This manuscript contained two proofs of the subsequence theorem, one by Martin Kruskal and one by Joseph Kruskal. At the time Joseph Kruskal did not know that Erdös and Szekeres had proven the subsequence theorem 15 years earlier. By the time Joseph Kruskal published the manuscript (which contained other things of interest) he had learned of the Erdös Szekeres paper and referenced it; however, he omitted his proof of the theorem and only sketched Martin Kruskal's proof, in the published version [3]

We use the following notation throughout.

Notation 0.1.2 If $n \in \mathbb{N}$ then $[n]$ is the set $\{1, \ldots, n\}$.

0.2 Proof by Erdös and Szekeres from 1935

Proof: We do this by induction on n. The base case of $n = 1$ is obvious.

We assume that the theorem is *true* for n and prove that it is *true* for $n+1$.

Let

$$x_1, \ldots, x_{(n+1)^2+1}$$

be a sequence of distinct reals. We view it as

$$x_1, \dots, x_{n^2+1}, x_{n^2+2}, \dots, x_{n^2+2n+2}$$

We now define $i_1, \dots, i_{2n+2} \in [n^2+2n+2]$ and $b_1, \dots, b_{2n+2} \in \{inc, dec\}$ as follows.

1. By the induction hypothesis

   $$x_1, \dots, x_{n^2+1}$$

   has an increasing or decreasing subsequence of length $n+1$. Let i_1 be such that x_{i_1} is the last element of some such subsequence. Let b_1 be *inc* if that subsequence is increasing, and *dec* if that subsequence is decreasing.

2. Remove x_{i_1} from the original sequence. (Note that what is left is now

   $$x_1, x_2, \dots, x_{i_1-1}, x_{i_1+1}, \dots, x_{n^2+1}.$$

)

3. Look at the first n^2+1 elements of the new sequence. Do the same process to obtain i_2, x_{i_2}, and b_2. Remove x_{i_2}.

4. Repeat this process until you have $i_1, \dots, i_{2n+2}$ and $b_1, \dots, b_{2n+2}$.

There are several cases.

Case 1: Assume $n+2$ of the b's are labeled *inc*. Look at just those $n+2$ elements, which we call

$$x_{j_1}, x_{j_2}, \dots, x_{j_{n+2}}.$$

These are all endpoints of subsequences of length $n+2$.

If there is an $a < b$ with $x_{j_a} < x_{j_b}$ then the increasing $n+1$-long subsequence that ends with x_{j_a}, together with x_{j_b}, forms an increasing $n+2$-long subsequence. If no such a, b exists then

$$x_{j_1}, x_{j_2}, \dots, x_{j_{n+2}}$$

is a decreasing $n+2$-long subsequence.

Case 2: Assume $n+2$ of the b's are labeled *dec*. Similar to Case 1.

Case 3: Assume $n+1$ of the b's are labeled *inc* and $n+1$ of the b's are labeled *dec*. Let $b_{j_1}, \ldots, b_{j_{n+1}}$ be labeled *dec* and let $b_{k_1}, \ldots, b_{k_{n+1}}$ be labeled *inc*. By the same reasoning as in Case 1 we have

$$x_{j_1} < \cdots < x_{j_{n+1}}$$

and

$$x_{k_1} > \cdots > x_{k_{n+1}}.$$

There are three subcases.

Subcase 3a: $x_{j_{n+1}} \neq x_{(n+1)^2+1}$ (the last point of the original sequence). There are two cases: $x_{j_{n+1}} < x^2_{n+1}$ or $x_{j_{n+1}} > x^2_{n+1}$. In the former case we get that

$$x_{j_1} < \cdots < x_{j_{n+1}} < x_{(n+1)^2+1}$$

is an increasing $(n+2)$-length subsequence. In the later we get that the decreasing $(n+1)$-length subsequence that ends with $x_{j_{n+1}}$ can be extended to a decreasing $(n+2)$-length subsequence by adding $x_{(n+1)^2+1}$ to the end.

Subcase 3b: $x_{k_{n+1}} \neq x_{(n+1)^2+1}$ (the last point of the original sequence). The reasoning is similar to Subcase 3a.

Subcase 3c: $x_{j_{n+1}} = x_{k_{n+1}} = x_{(n+1)^2+1}$ (the last point of the original sequence).

This cannot happen since once an element is picked to be one of the x's, it is removed.

∎

0.3 Proof by Martin Kruskal from 1948

Proof:

We do this by induction on n. The base case of $n = 1$ is obvious.

We assume that the statement is *false* at $n+1$ and show that it is *false* at n.

Since the statement is false at $n+1$ there is a sequence

$$x_1, x_2, \ldots, x_{(n+1)^2+1}$$

that has no increasing or decreasing subsequence of length $n+2$.

Let

$$A = \{x_i : (\forall j > i)[x_j < x_i]\}.$$

$$B = \{x_i : (\forall j > i)[x_j > x_i]\}.$$

Claim 1: $|A| \le n+1$ and $|B| \le n+1$.
Proof: We show $|A| \le n+1$. The proof that $|B| \le n+1$ is similar. By the definition of A, A is a decreasing sequence. Since the original sequence does not have a decreasing subsequence of length $n+2$, $|A| \le n+1$.
End of Proof of Claim 1
Claim 2: $|A \cap B| = 1$.
Proof: Note that $x_{(n+1)^2+1} \in A \cap B$ vacuously. Assume, by way of contradiction, that $i < (n+1)^2 + 1$ and $x_i \in A \cap B$. Either $x_i < x_{(n+1)^2+1}$ (so $x_i \notin A$) or $x_i > x_{(n+1)^2+1}$ (so $x_i \notin B$). Either way is a contradiction.
End of Proof of Claim 2

Combining Claims 1 and 2 we have $|A \cup B| \le 2n+1$. Remove $A \cup B$ from the sequence. The new sequence has length at least

$$(n+1)^2 + 1 - (2n+1) = n^2 + 1.$$

Claim 3: The new sequence has no increasing or decreasing $(n+1)$-subsequence

Assume, by way of contradiction, that there is an increasing $(n+1)$-subsequence in the new sequence (the case of decreasing is similar). Let it be

$x_{i_1} < \cdots < x_{i_{n+1}}$.

Note that $x_{i_{n+1}}$ was in the original sequence but was not removed. Hence there was some element x_j, with $j > i_{n+1}$ such that $x_{i_{n+1}} < x_j$. Hence

$$x_{i_1} < \cdots < x_{i_{n+1}} < x_j$$

is an increasing $(n+2)$-subsequence of the original sequence. This is a contradiction. Hence the new sequence cannot have an increasing $(n+1)$-subsequence.
End of Proof of Claim 3

Hence, assuming the theorem is false at $n+1$ we obtain that it is false at n.

▮

0.4 Proof by Joseph Kruskal from 1950

Proof:

We do this by induction on n. The base case of $n = 1$ is obvious.

We assume that the statement is *false* at $n + 1$ and show that it is *false* at n.
Since the statement is false at $n + 1$ there is a sequence

$$x_1, x_2, \ldots, x_{(n+1)^2+1}$$

that has no increasing or decreasing subsequence of length $n + 2$.

For $1 \le p \le n + 1$ let

$$A_p = \{x \mid \ x \text{ is the } p^{\text{th}} \text{ element of an increasing subsequence of length } n + 1\ \}.$$

Note that A_p could be empty.
Claim 1: There exists p such that $|A_p| \le n + 1$
Proof:

We first show that all of the A_p's are disjoint.
Assume, by way of contradiction, that there exists $1 \le p < q \le n + 1$ and $1 \le i \le (n + 1)^2 + 1$ such that $x_i \in A_p \cap A_q$.
Since $x_i \in A_p$ there is an increasing subsequence with x_i as the p^{th} element. Let it be

$$x_{j_1} < x_{j_2} < \cdots < x_{j_{p-1}} < x_i < x_{j_{p+1}} < \cdots < x_{j_{n+1}}$$

Since $x_i \in A_q$ there is an increasing subsequence with x_i as the q^{th} element. Let it be

$$x_{k_1} < x_{k_2} < \cdots < x_{k_{q-1}} < x_i < x_{k_{q+1}} < \cdots < x_{k_{n+1}}.$$

By splicing these two together we get the increasing subsequence

$$x_{k_1} < x_{k_2} < \cdots < x_{k_{q-1}} < x_i < x_{j_{p+1}} < \cdots < x_{j_{n+1}}.$$

This increasing subsequence is of size

$$(q - 1) + 1 + (n - p) = n + 1 + (q - p) \ge n + 2.$$

This contradicts the premise that there were no increasing subsequences of length $n + 2$.

Since

$$|A_1 \cup \cdots \cup A_{n+1}| \le (n+1)^2 + 1.$$

and all of the A_p's are disjoint, there exists p such that $|A_p| \le n+1$.
End of Proof of Claim 1

Let A_p be such that $|A_p| \le n+1$. Remove A_p and, if needed, more elements so that $n+1$ are removed. What is left has

$$(n+1)^2 + 1 - (n+1) = n^2 + 2n + 1 + 1 - n - 1 = n^2 + n + 1 \text{ elements .}$$

Now, for *this* sequence let

$$B_p = \{x \mid \; x \text{ is the } p^{\text{th}} \text{ element of a decreasing subsequence of length } n+1 \ \}.$$

By reasoning similar to Claim 1, there exists p such that $|B_p| \le n$.
Remove B_p. What is left has at most $n^2+n+1-n = n^2+1$ elements. Remove enough additional elements so that what is left has exactly n elements.
Claim 2: The final sequence has no increasing or decreasing subsequence of length $n+1$
Proof of Claim 2:

Assume, by way of contradiction, that there is an increasing subsequence of length $n+1$ (the proof for decreasing is similar). Let the sequence be

$x_{i_1} < \cdots < x_{i_{n+1}}$.

Look at x_{i_p}. It is the p^{th} element of an increasing sequence of length $n+1$. But x_{i_p} is not in A_p. Hence there must have been, in the original sequence, a increasing subsequence that contains $x_{i_1} < \cdots < x_{i_{n+1}}$ and one more element. This is an increasing subsequence of length $n+2$, a contradiction.
End of Proof of Claim 2:

Hence, assuming the theorem is false at $n+1$ we obtain that it is false at n.

∎

0.5 Proof Attributed to Erdos and Szekeres from Proofs from the Book [1]

Proof: Let $x_1, \ldots, x_{n^2+1}$ be any sequence of distinct reals. Assume, by way of contradiction, that there are no increasing or decreasing $(n+1)$-subsequences.

Let f be the function with domain $[n^2+1]$ and co-domain $[n]$ that is defined by

$f(i) =$ length of longest increasing subsequence that ends with x_i.

Since f has domain of size n^2+1 and range of size n, there exists $i_1, \ldots, i_n, i_{n+1}$ and a such that

$$f(x_{i_1}) = f(x_{i_2}) = \cdots = f(x_{i_{n+1}}) = a.$$

Let $1 \le j \le n$. Note that $x_{i_j} > x_{i_{j+1}}$ since otherwise there would be an increasing $a+1$ subsequence that ends with $x_{i_{j+1}}$ (take the increasing a-subsequence that ends with x_{i_j} and add $x_{i_{j+1}}$ to it). Hence

$$x_{i_1} > x_{i_2} > \cdots > x_{i_{n+1}}$$

is a decreasing $(n+1)$-subsequence, which is a contradiction. ▮

0.6 Proof that I do not know where it came from

Proof:

Let $x_1, \ldots, x_{n^2+1}$ be any sequence of distinct reals. Let f be the function with domain $[n^2+1]$ and co-domain $[n+1] \times [n+1]$ that is defined by $f(i) = (a, b)$ where

1. a is length of longest increasing subsequence that ends with x_i.
2. b is length of longest decreasing subsequence that ends with x_i.

Claim 1: f is injective

Assume, by way of contradiction, that there is an $i < j$ such that

$$f(i) = f(j) = (a, b).$$

If $x_i < x_j$ then x_j is the last point in a sequence of length $a+1$ (just take the sequence of length a that ends with x_i and add x_j to it). Hence $f(j)$ would have first coordinate at least $a+1$. If $x_i > x_j$ then, by similar reasoning, $f(j)$ would have second coordinate at least $b+1$. Hence $f(j) \neq (a,b)$.

End of Proof of Claim 1

Claim 2: There exists i such that $f(i) = (a,b)$ where either $a \geq n+1$ or $b \geq n+1$

The domain of f has n^2+1 elements. Hence, since f is injective, the image of f must have n^2+1 elements.

Assume, by way of contradiction, that the image of f is contained in $[n] \times [n]$. This is of size n^2. This is a contradiction.

End of Proof of Claim 2

Since there is an i such that $f(i) = (a,b)$ with either $a \geq n+1$ or $b \geq n+1$, there is an increasing or decreasing $(n+1)$-subsequence.

∎

Bibliography

[1] M. Aigner and G. Ziegler. *Proofs from the book*. Springer, 2003 (third edition).

[2] P. Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio Math*, 2:463–470, 1935.

[3] J. Kruskal. Monotonic subsequences. *Proceedings of the American Mathematical Society*, 4:264–274, 1953.