Explanation to Wikipedia's Variation of Parameter

Vincent Tam

March 22, 2014

Wikipedia's explanation is too concise. [wiki12]

First, you have

$$y^{(n)}(x) + \sum_{k=0}^{n-1} a_k(x) y^{(k)}(x) = b(x). \tag{i}$$

Here, y means the general solution of (i).

I just use the notations in Wikipedia: For i between 1 and n, inclusive, y_i is the solution of the homogenous equation

$$y_i^{(n)}(x) + \sum_{k=0}^{n-1} a_k(x) y_i^{(k)}(x) = 0. \tag{ii}$$

Unlike Wikipedia, in equation (ii), the i under the summation sign is changed to k so as to avoid confusion of symbols. In other words, **we use i to represent $c_i(x)$ and $y_i(x)$, while using k to represent the coefficient of the terms in** (i) **(i.e. $a_k(x)$) and the derivatives of $y_i(x)$ of different orders. (i.e. $y_i^{(k)}(x)$**
Let y_p be the particular solution for the non-homogeneous ODE.

We make an educated guess like this:

$$\boxed{y_p(x) = \sum_{i=1}^{n} c_i(x) y_i(x)} \tag{iii}$$

where $c_i(x)$ is a differentiable function for each i.
By doing so, we have created n unknowns (that is, $c_i(x), i = 1, \ldots n$), and they are also known as *freedom*. We may make use of the freedom like this:

$$y_p'(x) = \left[\sum_{i=1}^{n} c_i'(x) y_i(x)\right] + \left[\sum_{i=1}^{n} c_i(x) y_i'(x)\right] \tag{iv}$$

Continuing the differentiation gives complicated expression, so we use our freedom to give some restrictions on $c_i(x), i = 1, \ldots n$. i.e. We set

$$\boxed{\sum_{i=1}^{n} c_i'(x) y_i(x) = 0} \tag{v}$$

Why is there such a condition? It is because we want to simplify (iv) by ruling out *something*. (i.e. setting something to *zero*) What is *something*?

1. $y_i(x) = 0$ or $y_i'(x) = 0$ or $c_i(x) = 0$ for each $i = 1, \ldots n$
   They are clearly unreasonable restrictions.

2. $c_i'(x) = 0$ for each $i = 1, \ldots n$
   Then $c_i(x)$ is a constant function for each $i = 1, \ldots n$. In other words, $y_p(x) = \sum_{i=1}^{n} c_i(x) y_i(x)$ is just linear combinations of $y_i(x)$. Remembering that $y_i(x)$ are denoted as the solution of (ii), we conclude that this restriction can't be accepted.

Being too "strong", the above guesses fails. Thus, what we need to rule out is *either* one of the middle bracket in (iv). Why *must* it be the first one?

Otherwise assume that

$$\sum_{i=1}^{n} c_i(x) y_i'(x) = 0, \tag{vi}$$

which leads us to

$$y_p'(x) = \sum_{i=1}^{n} c_i'(x) y_i(x) \tag{vii}$$

Continue the above process of differentiating $y_p(x)$.

$$y_p''(x) = \left[\sum_{i=1}^{n} c_i''(x) y_i(x) \right] + \left[\sum_{i=1}^{n} c_i'(x) y_i'(x) \right] \tag{viii}$$

Making assumptions similar to (vi), we have

$$\sum_{i=1}^{n} c_i'(x) y_i'(x) = 0; \text{ and} \tag{ix}$$

$$y_p''(x) = \sum_{i=1}^{n} c_i''(x) y_i(x) \tag{x}$$

From (vi) and (ix), we know that if we carry out the process repeatedly, we get a system of "linear" equations with "unknowns " $y_i'(x)$ (which is actually known since $y_i(x)$ have been found out from (ii).) The most insensible thing of this "system" is the "coefficient matrix" with entries $c_i^{(k)}$ on the k-th row and the i-th column. For the sake of clarity, the "system" is illustrated below.

$$\begin{cases} c_1(x)y_1'(x) + c_2(x)y_2'(x) & + \ldots + c_n(x)y_n'(x) & = 0 \\ c_1'(x)y_1'(x) + c_2'(x)y_2'(x) & + \ldots + c_n'(x)y_n'(x) & = 0 \\ \qquad \vdots & & \\ c_1^{(k)}(x)y_1'(x) + c_2^{(k)}(x)y_2'(x) & + \ldots + c_n^{(k)}(x)y_n'(x) & = 0 \end{cases} \tag{*}$$

We don't care about the range of k such that the R.H.S. of the last equation in the above system is zero, as well as the missing equation which can be yielded after substituting $y_p^{(k)}(x)$ (for example, equations (vii) and (x)) into (i)

Therefore, equations (vi) to (*) are all *wrong*. Instead, we should have

$$\boxed{y_p'(x) = \sum_{i=1}^{n} c_i(x)y_i'(x)} \tag{xi}$$

$$y_p''(x) = \left[\sum_{i=1}^{n} c_i'(x)y_i'(x)\right] + \left[\sum_{i=1}^{n} c_i(x)y_i''(x)\right] \tag{xii}$$

Again, we set

$$\boxed{\sum_{i=1}^{n} c_i'(x)y_i'(x) = 0} \tag{xiii}$$

so that

$$\boxed{y_p''(x) = \sum_{i=1}^{n} c_i(x)y_i''(x).} \tag{xiv}$$

For each $1 \leq k \leq n-1$, we have

$$\boxed{\sum_{i=1}^{n} c_i'(x)y_i^{(k-1)}(x) = 0} \tag{xv}$$

and

$$\boxed{y_p^{(k)}(x) = \sum_{i=1}^{n} c_i(x)y_i^{(k)}(x).} \tag{xvi}$$

Why *can't* we apply (xv) and (xvi) to the case of $k = n$?
Differentiating (xvi) with respect to x, we get

$$y_p^{(n)}(x) = \left[\sum_{i=1}^{n} c_i'(x)y_i^{(n-1)}(x)\right] + \left[\sum_{i=1}^{n} c_i(x)y_i^{(n)}(x)\right] \tag{xvii}$$

Recall from (i) that we have

$$y_p^{(n)}(x) + \sum_{k=0}^{n-1} a_k(x) y_p^{(k)}(x) = b(x). \tag{xviii}$$

Substitute (xvi) and (xvii) into (xviii),

$$\left[\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)}(x)\right] + \left[\sum_{i=1}^{n} c_i(x) y_i^{(n)}(x)\right] + \left\{\sum_{k=0}^{n-1} a_k(x) \left[\sum_{i=1}^{n} c_i(x) y_i^{(k)}(x)\right]\right\} = b(x).$$

$$\left[\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)}(x)\right] + \left[\sum_{i=1}^{n} c_i(x) y_i^{(n)}(x)\right] + \left\{\sum_{i=1}^{n} c_i(x) \left[\sum_{k=0}^{n-1} a_k(x) y_i^{(k)}(x)\right]\right\} = b(x).$$

$$\left[\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)}(x)\right] + \left\{\sum_{i=1}^{n} c_i(x) \left[y_i^{(n)} + \sum_{k=0}^{n-1} a_k(x) y_i^{(k)}(x)\right]\right\} = b(x)$$

$$\left[\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)}(x)\right] + \left[\sum_{i=1}^{n} c_i(x)(0)\right] = b(x)$$

Thus, we have

$$\boxed{\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)}(x) = 0} \tag{xix}$$

From equations (xv) and (xix),

$$\begin{cases} c_1'(x)y_1(x) + c_2'(x)y_2(x) & + \ldots + c_n'(x)y_n(x) & = 0 \\ c_1'(x)y_1'(x) + c_2'(x)y_2'(x) & + \ldots + c_n'(x)y_n'(x) & = 0 \\ \vdots & & \\ c_1'(x)y_1^{(n-1)}(x) + c_2'(x)y_2^{(n-1)}(x) & + \ldots + c_n'(x)y_n^{(n-1)}(x) & = 0 \\ c_1'(x)y_1^{(n)}(x) + c_2'(x)y_2^{(n)}(x) & + \ldots + c_n'(x)y_n^{(n)}(x) & = b(x) \end{cases} \tag{$\heartsuit$}$$

By Cramer's Rule,

$$c_i'(x) = \frac{W_i(x)}{W(x)},$$

where $W(x)$ is the Wronskian of $y_i(x)$ and $W_i(x)$ is the determinant of replacing the i-th column of $W(x)$ by $(0, 0, \ldots, 0, b(x))$. Hence,

$$c_i(x) = \int \frac{W_i(x)}{W(x)} \, \mathrm{d}x. \tag{xx}$$

Substitute (xx) into (iii),

$$\boxed{y_p(x) = \sum_{i=1}^{n} y_i(x) \int \frac{W_i(x)}{W(x)} \, \mathrm{d}x}$$

References

[wiki12] Variation of Parameters. (n.d.). In *Wikipedia*. Retrieved October 26, 2013, from https://en.wikipedia.org/wiki/Variation_of_parameters