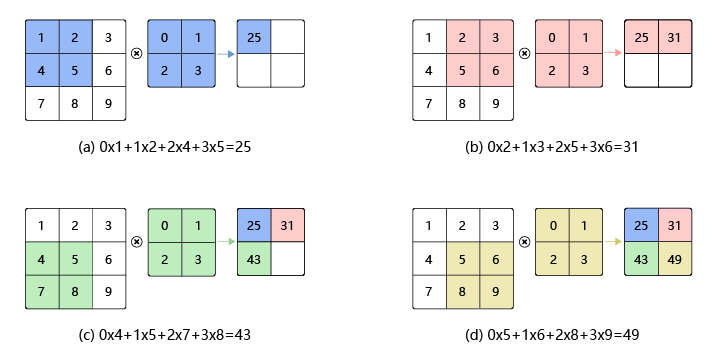
实验二 基于SIMD指令集的程序设计与性能分析

一、实验内容

使用**SIMD指令集**(可以结合其他优化方式)，对卷积算法进行优化，分析实验结果是否符合预期，探讨优化结果符合预期的原因或导致优化效果未达预期的因素。

二、实验说明

1. 卷积算法是一种数学运算，主要在两个函数或数据集合上执行，本实验考虑的是两个矩阵之间进行的运算，代码可参考附件”convolution.c”。通常，其中一个是“输入矩阵”，规模较大。另一个是“核”，规模较小。卷积执行过程中，从输入矩阵的左上角开始。核覆盖输入矩阵的一部分，这部分与核的大小相匹配。然后，将核的每个元素与其下方输入矩阵相对应位置的元素相乘，并将这些乘积求和，结果作为卷积操作的结果的一部分存储在一个新的“输出矩阵”中。然后滑动”核”，覆盖新的部分元素，重复这个过程并将结果存储在输出矩阵的新的位置。如下图所示：



一个输入矩阵为3\*3，核为2\*2，步长为1的卷积示意图

卷积计算可参考：<https://www.cnblogs.com/MrSaver/p/10356293.html>

卷积代码可参考附件” convolution.c”，**实验中，可以修改代码但不可偏离主题。例如可以修改核大小或者数据规模，提高SIMD指令加速效果。**

1. SIMD指令集

SIMD指令集是一种允许同时加载多个位宽的数据到一个寄存器中并进行多种操作计算的CPU扩展指令。关于SIMD的基础知识可以参考这篇博客 [知乎-玩转SIMD指令编程](https://zhuanlan.zhihu.com/p/591900754)。在正式开始实验前，需要先查看你的电脑支持哪些CPU指令集（可以使用CPU-Z等软件），目前主流的X86\_64处理器已经支持了SSE4.2和AVX2等常用SIMD指令集。

虽然SIMD是汇编指令，但是我们不需要直接使用汇编开发，在[intel的SIMD文档](https://www.intel.com/content/www/us/en/docs/intrinsics-guide/index.html#expand=914)中，给出了一个SIMD的C接口规范，并提供了详细的说明和底层原理伪代码描述，请自行查找资料学习SIMD编程方法，然后设计程序对上述提到的卷积运算进行加速。

三、考察重点

**1. SIMD编程**

**2. 实验设计、数据统计方法与分析结论**

四、注意事项

1. SIMD指令集的头文件根据编译器的实现不同而略有差异，有的代码的写法是依赖编译器版本的，但是文档中提到的方法都会被实现。
2. “其他优化方式”包括但不限于利用程序的时间、空间局部性进行程序算法的优化，如果你用到了这些方法，请在报告中说明这些方法的原理和效果。
3. 关于矩阵的输入规模，请自行调整，方便测试即可。一般来说，卷积核的长度是奇数的，为了降低难度，使用偶数长度的卷积核即可。另外卷积核规模不宜过大。
4. 矩阵中单个数据的位宽不作限制，可以灵活调整，多做尝试。
5. 在报告中展示加速思路、方法、数据规模和对应的加速效果。程序实际执行时间还取决于编译器种类、优化选项、操作系统、CPU型号等因素，本实验需要控制这些无关因素不变，对比原始卷积运算和优化后的加速效果。

五、实验报告要求

1. 4人一组完成，提交一份实验报告（尽量凑齐4个人一组）；

2. 提交形式为.zip压缩文件，命名方式为“张三-李四-王五-赵六-实验二报告.zip”（命名方式不正确会扣分）。压缩包中需中包括修改后的代码文件和实验报告文件，实验报告格式为word或pdf，并在报告开头写上全组成员名字学号。发送到山大网盘：

<https://icloud.qd.sdu.edu.cn:7777/link/683BA15450B0E7897BA882B5552A8A91>

3. 提交报告的截止时间为4月18日（周四晚上11点59分59秒）。晚交的报告将会被扣分。

4. 所有雷同的实验报告（包括和之前高年级提交的实验报告雷同），本次实验不得分。同一学生本学期出现2次及以上雷同的实验报告，本学期实验为0分。