# [整数，小数及常用的正则表达式](http://blog.csdn.net/shenlan211314/article/details/5315450)

匹配中文字符的正则表达式： [/u4e00-/u9fa5]  
评注：匹配中文还真是个头疼的事，有了这个表达式就好办了

匹配双字节字符(包括汉字在内)：[^/x00-/xff]  
评注：可以用来计算字符串的长度（一个双字节字符长度计2，ASCII字符计1）

匹配空白行的正则表达式：/n/s\*/r  
评注：可以用来删除空白行

匹配HTML标记的正则表达式：<(/S\*?)[^>]\*>.\*?<//1>|<.\*? />  
评注：网上流传的版本太糟糕，上面这个也仅仅能匹配部分，对于复杂的嵌套标记依旧无能为力

匹配首尾空白字符的正则表达式：^/s\*|/s\*$  
评注：可以用来删除行首行尾的空白字符(包括空格、制表符、换页符等等)，非常有用的表达式

匹配Email地址的正则表达式：/w+([-+.]/w+)\*@/w+([-.]/w+)\*/./w+([-.]/w+)\*  
评注：表单验证时很实用

匹配网址URL的正则表达式：[a-zA-z]+://[^/s]\*  
评注：网上流传的版本功能很有限，上面这个基本可以满足需求

匹配帐号是否合法(字母开头，允许5-16字节，允许字母数字下划线)：^[a-zA-Z][a-zA-Z0-9\_]{4,15}$  
评注：表单验证时很实用

匹配国内电话号码：/d{3}-/d{8}|/d{4}-/d{7}  
评注：匹配形式如 0511-4405222 或 021-87888822

匹配腾讯QQ号：[1-9][0-9]{4,}  
评注：腾讯QQ号从10000开始

匹配中国邮政编码：[1-9]/d{5}(?!/d)  
评注：中国邮政编码为6位数字

匹配身份证：/d{15}|/d{18}  
评注：中国的身份证为15位或18位

匹配ip地址：/d+/./d+/./d+/./d+  
评注：提取ip地址时有用

匹配特定数字：  
^[1-9]/d\*$　 　 //匹配正整数  
^-[1-9]/d\*$ 　 //匹配负整数  
^-?[1-9]/d\*$　　 //匹配整数  
^[1-9]/d\*|0$　 //匹配非负整数（正整数 + 0）  
^-[1-9]/d\*|0$　　 //匹配非正整数（负整数 + 0）  
^[1-9]/d\*/./d\*|0/./d\*[1-9]/d\*$　　 //匹配正浮点数  
^-([1-9]/d\*/./d\*|0/./d\*[1-9]/d\*)$　 //匹配负浮点数  
^-?([1-9]/d\*/./d\*|0/./d\*[1-9]/d\*|0?/.0+|0)$　 //匹配浮点数  
^[1-9]/d\*/./d\*|0/./d\*[1-9]/d\*|0?/.0+|0$　　 //匹配非负浮点数（正浮点数 + 0）  
^(-([1-9]/d\*/./d\*|0/./d\*[1-9]/d\*))|0?/.0+|0$　　//匹配非正浮点数（负浮点数 + 0）  
评注：处理大量数据时有用，具体应用时注意修正

匹配特定字符串：  
^[A-Za-z]+$　　//匹配由26个英文字母组成的字符串  
^[A-Z]+$　　//匹配由26个英文字母的大写组成的字符串  
^[a-z]+$　　//匹配由26个英文字母的小写组成的字符串  
^[A-Za-z0-9]+$　　//匹配由数字和26个英文字母组成的字符串  
^/w+$　　//匹配由数字、26个英文字母或者下划线组成的字符串  
评注：最基本也是最常用的一些表达式

在Perl:

从ISO 8859-1到UTF-8: s/([/x80-/xFF])/chr(0xC0|ord($1)>>6).chr(0x80|ord($1)&0x3F)/eg;

从UTF-8到ISO 8859-1: s/([/xC2/xC3])([/x80-/xBF])/chr(ord($1)<<6&0xC0|ord($2)&0x3F)/eg;