Subject: Re: matching lists
Posted by Andy Loughe on Fri, 10 Mar 2000 08:00:00 GMT
View Forum Message <> Reply to Message

; RESTRICTIONS:
;   if the indices are not unique some matching elements may be
skipped...
;   or is it worse than that?

Maybe worse than that...
   Isn't a significant restriction that the lists must be numeric?
   Or does it actually work for alpha-numeric lists?

Mark Fardal wrote:
>
> Hi,
>
> I have been looking at properties of particles in a simulation, and
> sometimes need to match up the particles in two different subsets. I
> typically have (quantity A, index #) for one set of particles, and
> (quantity B, index #) for another set, and want to compare quantities
> A and B for the particles that are in both sets.
>
> As of late last night I could not think of a good way to do this;
> WHERE inside a for-loop would be very slow. Maybe I'm missing
> something easy, but in any case here's a solution inspired by the
> recently submitted SETINTERSECTION function. Hope somebody finds
> it useful.
>
> Mark Fardal
> UMass
>
> ;+
> ; NAME:
> ;   LISTMATCH
> ;
> ; PURPOSE:
> ;   find the indices where a matches b
> ;   works only for SETS OF UNIQUE INTEGERS, e.g., array indices
> ;
> ;   for example: suppose you have a list of people with their ages and
> ;   social security numbers (AGE, AGE_SS), and a partially overlapping
> ;   list of people with their incomes and s.s. numbers (INCOME,
> ;   INCOME_SS). And you want to correlate ages with incomes in the
> ;   overlapping subset. Call
> ;   LISTMATCH, AGE_SS, INCOME_SS, AGE_IND, INCOME_IND

```
> ;   then AGE[AGE_IND] and INCOME[INCOME_IND] will be the desired
> ;   pair of variables.
> ;
> ; AUTHOR:
> ;   Mark Fardal
> ;   UMass (fardal@weka.astro.umass.edu)
> ;
> ; CALLING SEQUENCE:
> ;   LISTMATCH, a, b, a_ind, b_ind
> ;
> ; INPUTS:
> ;   a and b are sets of unique integers (no duplicate elements)
> ;
> ; OUTPUTS:
> ;   a_ind, b_ind are the indices indicating which elements of a and b
> ;   are in common
> ;
> ; RESTRICTIONS:
> ;   if the indices are not unique some matching elements may be skipped...
> ;   or is it worse than that?
> ; EXAMPLE:
> ;   a = [2,4,6,8]
> ;   b = [6,1,3,2]
> ;   listmatch, a, b, a_ind, b_ind
> ;    print, a[a_ind]
> ;       2     6
> ;    print, b[b_ind]
> ;       2     6
> ;
> ;
> ; MODIFICATION HISTORY:
> ;   none
> ; BUGS:
> ;   tell me about them
> ; ACKNOWLEDGEMENTS:
> ;   trivial modification of SETINTERSECTION from RSI
> ;
> pro listmatch, a, b, a_ind, b_ind
> minab = min(a, MAX=maxa) > min(b, MAX=maxb) ;Only need intersection of ranges
> maxab = maxa < maxb
> ;If either set is empty, or their ranges don't intersect:
> ;  result = NULL (which is denoted by integer = -1)
> if maxab lt minab or maxab lt 0 then begin
>   a_ind = -1
>   b_ind = -1
>   return
> endif
>
```

```
> ha = histogram(a, MIN=minab, MAX=maxab, reverse_indices=reva)
> hb = histogram(b, MIN=minab, MAX=maxab, reverse_indices=revb)
>
> r = where((ha ne 0) and (hb ne 0), count)
> if count gt 0 then begin
>   a_ind = reva[reva[r]]
>   b_ind = revb[revb[r]]
>   return
> endif else begin
>   a_ind = -1
>   b_ind = -1
>   return
> endelse
>
> end
>
>


--
--------
Andrew F. Loughe    email:loughe@fsl.noaa.gov    phone:(303)497-6211
```