

Hidden Markov model example with MATLAB

1. Fixing a point on the variety:

S=[0.8,0.2; 0.1 0.9]

S =

0.800000000000000 0.200000000000000
0.100000000000000 0.900000000000000

T=[0.25 0.25 0.25 0.25; 0.125, 0.375, 0.375, 0.125]

T =

0.250000000000000 0.250000000000000 0.250000000000000 0.250000000000000
0.125000000000000 0.375000000000000 0.375000000000000 0.125000000000000

2. Fixing a sequence sigma:

DNaseq=hmmgenerate(100,S,T,'Statenames',{'exon','intron'},'Symbols',{'A','C','G','T'})

DNaseq =

Columns 1 through 14

'G' 'C' 'C' 'C' 'C' 'A' 'G' 'A' 'C' 'T' 'C' 'T' 'A' 'C' 'C'

Columns 15 through 28

'T' 'C' 'T' 'C' 'C' 'G' 'C' 'G' 'G' 'G' 'C' 'G' 'G' 'G'

Columns 29 through 42

'C' 'A' 'T' 'T' 'T' 'G' 'A' 'G' 'G' 'C' 'C' 'A' 'G' 'G'

Columns 43 through 56

'C' 'G' 'C' 'T' 'T' 'A' 'C' 'C' 'G' 'T' 'G' 'G' 'G' 'G'

Columns 57 through 70

'C' 'C' 'C' 'A' 'A' 'C' 'T' 'C' 'A' 'C' 'G' 'G' 'G' 'G'

Columns 71 through 84

'C' 'G' 'A' 'C' 'T' 'C' 'G' 'C' 'G' 'A' 'T' 'A' 'T' 'G'

Columns 85 through 98

'A' 'A' 'G' 'A' 'C' 'A' 'C' 'T' 'G' 'G' 'A' 'A' 'T'

Columns 99 through 100

'T' 'G'

3. Evaluating the sigma coordinate of the point (answer in logpseq):

[PSTATES,logpseq]=hmmdecode(DNAseq,S,T,'Symbols',{'A','C','G','T'})

PSTATES =

Columns 1 through 5

0.60269417895039 0.39545352949766 0.29488836174298 0.26053316770887 0.27856407170065
0.39730582104961 0.60454647050234 0.70511163825702 0.73946683229113 0.72143592829935

Columns 6 through 10

0.35623636603883 0.32096815699868 0.34996540284456 0.45489604297638 0.45112973926224
0.64376363396117 0.67903184300132 0.65003459715544 0.54510395702363 0.54887026073776

Columns 11 through 15

0.53432855985292 0.51984806975245 0.40096518427134 0.36782470824217 0.40709154535795
0.46567144014708 0.48015193024755 0.59903481572866 0.63217529175784 0.59290845464205

Columns 16 through 20

0.33710609259936 0.33617816108945 0.21616194703670 0.15201458312126 0.11792439195811
0.66289390740064 0.66382183891055 0.78383805296330 0.84798541687874 0.88207560804189

Columns 21 through 25

0.10017412996019 0.09162142979974 0.08882484784079 0.09065909277127 0.09786222981322
0.89982587003981 0.90837857020026 0.91117515215921 0.90934090722873 0.90213777018678

Columns 26 through 30

0.11333266410970 0.14329540374668 0.19980688443524 0.30560624959528 0.50326514854121

0.88666733589030	0.85670459625332	0.800193111556476	0.69439375040472	0.49673485145879
Columns 31 through 35				
0.57589005090898	0.55719966136935	0.43851629906462	0.41036520455833	0.26185456395850
0.42410994909102	0.44280033863065	0.56148370093538	0.58963479544167	0.73814543604150
Columns 36 through 40				
0.18461164956697	0.1475538392618	0.13577503157590	0.14453040311807	0.17734449329225
0.81538835043303	0.85244461607382	0.86422496842410	0.85546959688193	0.822265550670775
Columns 41 through 45				
0.24742106695425	0.20205194562013	0.19663959662293	0.22900619381851	0.31217543941179
0.75257893304575	0.79794805437987	0.80336040337707	0.77099380618149	0.68782456058821
Columns 46 through 50				
0.47961305365337	0.52795684350717	0.47965213997672	0.31227175928006	0.22915878536248
0.52038694634663	0.47204315649283	0.52034786002327	0.68772824071994	0.77084121463752
Columns 51 through 55				
0.19687014062384	0.20241348944309	0.24801936979752	0.17818891613869	0.14577936197373
0.80312985937616	0.79758651055691	0.75198063020247	0.82181108386131	0.85422063802626
Columns 56 through 60				
0.13774972003153	0.15086901534009	0.19041620244025	0.27230431614135	0.42948357362311
0.86225027996847	0.84913098465991	0.80958379755975	0.72769568385865	0.57051642637689
Columns 61 through 65				
0.46575414263798	0.39795593022857	0.40851150403450	0.30705849521450	0.27404712374403
0.53424585736202	0.60204406977143	0.59148849596550	0.69294150478550	0.72595287625597
Columns 66 through 70				
0.29619424253139	0.20150177781691	0.15405690611393	0.13476871476101	0.13587600295185
0.70380575746861	0.79849822218309	0.84594309388607	0.86523128523899	0.86412399704815
Columns 71 through 75				
0.15782432236325	0.20944525866311	0.31151009336259	0.29627346964088	0.34094915399206
0.84217567763675	0.79055474133689	0.68848990663741	0.70372653035912	0.65905084600794
Columns 76 through 80				

0.27240670477681	0.26117394975167	0.302731042225174	0.41379976473540	0.63907205553536
0.72759329522319	0.73882605024833	0.69726895774826	0.58620023526460	0.36092794446464

Columns 81 through 85

0.74644219703297	0.78576061206641	0.77528227904407	0.71014225763415	0.75156624891788
0.25355780296703	0.21423938793359	0.22471772095593	0.28985774236585	0.24843375108212

Columns 86 through 90

0.73503903600748	0.65288727005171	0.67590397916883	0.62509384745648	0.67642138425783
0.26496096399252	0.34711272994829	0.32409602083117	0.37490615254352	0.32357861574217

Columns 91 through 95

0.65399384655019	0.73685545344791	0.75434445224422	0.71458073523743	0.59910257667437
0.34600615344981	0.26314454655209	0.24565554775578	0.28541926476257	0.40089742332563

Columns 96 through 100

0.58946171777194	0.68177886054322	0.69697068436274	0.64209053600384	0.49165834658560
0.41053828222806	0.31822113945678	0.30302931563726	0.35790946399616	0.50834165341440

logpseq =

-1.341061974974420e+02

4. Tropical evaluation of the coordinate sigma of the point:

STATES=hmmviterbi(DNAseq,S,T,'Statenames',{'exon','intron'},'Symbols',{'A','C','G','T'})

STATES =

Columns 1 through 8

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 9 through 16

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 17 through 24

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 25 through 32

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 33 through 40

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 41 through 48

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 49 through 56

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 57 through 64

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 65 through 72

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'intron'

Columns 73 through 80

'intron' 'intron' 'intron' 'intron' 'intron' 'intron' 'exon'

Columns 81 through 89

'exon' 'exon' 'exon' 'exon' 'exon' 'exon' 'exon'

Columns 90 through 98

'exon' 'exon' 'exon' 'exon' 'exon' 'exon' 'exon'

Columns 99 through 100

'exon' 'exon'

6. Mechanics of the calculation: the forward and backward variables.

[PSTATES, logpseq, FORWARD, BACKWARD, SCALE]=hmmdecode(DNaseq, S, T, 'Symbols', {'A', 'C', 'G', 'T'})

PSTATES =

Columns 1 through 5

hummex1 Tue Sep 07 12:27:18 2004 6

0.60269417895039 0.39545352949766 0.29488836174298 0.26053316770887 0.27856407170065
0.39730582104961 0.60454647050234 0.705111163825702 0.73946683229113 0.72143592829935

Columns 6 through 10

0.35623636603883 0.32096815699868 0.34996540284456 0.45489604297638 0.45112973926224
0.64376363396117 0.67903184300132 0.65003459715544 0.54510395702363 0.54887026073776

Columns 11 through 15

0.5343285985292 0.51984806975245 0.40096518427134 0.36782470824217 0.40709154535795
0.46567144014708 0.48015193024755 0.59903481572866 0.63217529175784 0.59290845464205

Columns 16 through 20

0.33710609259936 0.33617816108945 0.21616194703670 0.15201458312126 0.11792439195811
0.66289390740064 0.66382183891055 0.78383805296330 0.84798541687874 0.88207560804189

Columns 21 through 25

0.10017412996019 0.09162142979974 0.08882484784079 0.09065909277127 0.09786222981322
0.89982587003981 0.90837857020026 0.91117515215921 0.90934090722873 0.90213777018678

Columns 26 through 30

0.1133266410970 0.14329540374668 0.19980688443524 0.30560624959528 0.50326514854121
0.88666733589030 0.85670459625332 0.80019311556476 0.69439375040472 0.49673485145879

Columns 31 through 35

0.57589005090898 0.55719966136935 0.43851629906462 0.41036520455833 0.26185456395850
0.42410994909102 0.44280033863065 0.56148370093538 0.58963479544167 0.73814543604150

Columns 36 through 40

0.18461164956697 0.14755538392618 0.13577503157590 0.14453040311807 0.17734449329225
0.81538835043303 0.85244461607382 0.86422496842410 0.85546959688193 0.82265550670775

Columns 41 through 45

0.24742106695425 0.20205194562013 0.19663959662293 0.22900619381851 0.31217543941179
0.75257893304575 0.79794805437987 0.80336040337707 0.77099380618149 0.68782456058821

Columns 46 through 50

0.47961305365337 0.52795684350717 0.47965213997672 0.31227175928006 0.22915878536248
0.52038694634663 0.47204315649283 0.52034786002327 0.68772824071994 0.77084121463752

Columns 51 through 55

0.19687014062384 0.20241348944309 0.24801936979752 0.17818891613869 0.14577936197373
0.80312985937616 0.79758651055691 0.75198063020247 0.82181108386131 0.85422063802626

Columns 56 through 60

0.13774972003153 0.15086901534009 0.19041620244025 0.27230431614135 0.42948357362311
0.86225027996847 0.84913098465991 0.80958379755975 0.72769568385865 0.57051642637689

Columns 61 through 65

0.46575414263798 0.39795593022857 0.40851150403450 0.30705849521450 0.27404712374403
0.53424585736202 0.60204406977143 0.59148849596550 0.69294150478550 0.72595287625597

Columns 66 through 70

0.29619424253139 0.20150177781691 0.15405690611393 0.13476871476101 0.13587600295185
0.70380575746861 0.79849822218309 0.84594309388607 0.86523128523899 0.86412399704815

Columns 71 through 75

0.15782432236325 0.20944525866311 0.31151009336259 0.29627346964088 0.34094915399206
0.84217567763675 0.79055474133689 0.68848990663741 0.70372653035912 0.65905084600794

Columns 76 through 80

0.27240670477681 0.26117394975167 0.30273104225174 0.41379976473540 0.63907205553536
0.72759329522319 0.73882605024833 0.69726895774826 0.58620023526460 0.36092794446464

Columns 81 through 85

0.74644219703297 0.78576061206641 0.77528227904407 0.71014225763415 0.75156624891788
0.25355780296703 0.21423938793359 0.22471772095593 0.28985774236585 0.24843375108212

Columns 86 through 90

0.73503903600748 0.65288727005171 0.67590397916883 0.62509384745648 0.67642138425783
0.26496096399252 0.34711272994829 0.32409602083117 0.37490615254352 0.32357861574217

Columns 91 through 95

0.65399384655019 0.73685545344791 0.75434445224422 0.71458073523743 0.59910257667437
0.34600615344981 0.26314454655209 0.24565554775578 0.28541926476257 0.40089742332563

Columns 96 through 100

0.58946171777194 0.68177886054322 0.69697068436274 0.64209053600384 0.49165834658560
0.41053828222806 0.31822113945678 0.30302931563726 0.35790946399616 0.50834165341440

logpseq =

-1.341061974974420e+02

FORWARD =

Columns 1 through 5

1.000000000000000	0.727272727272727	0.50950570342205	0.35909702496636	0.26532029102485
0	0.272727272727272	0.49049429657795	0.64090297503364	0.73467970897515

Columns 6 through 10

0.21053439308468	0.39663174036656	0.28801731089915	0.22354986071493	0.40825783420085
0.78946560691532	0.60336825963344	0.71198268910085	0.77645013928507	0.59174216579915

Columns 11 through 15

0.29514008412125	0.46930738462377	0.59994486368055	0.41931718767270	0.30195692126096
0.70485991587875	0.53069261537623	0.40005513631945	0.58068281232730	0.69804307873904

Columns 16 through 20

0.47487723787111	0.33682538984935	0.50274496208593	0.35471549934831	0.26270012530185
0.52512276212889	0.66317461015065	0.49725503791407	0.64528450065169	0.73729987469815

Columns 21 through 25

0.20904167863522	0.17890967417912	0.16234666053411	0.15334907658032	0.14849260335693
0.79095832136478	0.82109032582088	0.83765333946589	0.84665092341968	0.851507396664307

Columns 26 through 30

0.14588039891348	0.14447796935928	0.14372579475250	0.14332259316537	0.14310652025475
0.85411960108652	0.85552203064072	0.85627420524750	0.85667740683463	0.85689347974525

Columns 31 through 35

0.33357574831686	0.50019087750696	0.62081674382283	0.43365424592481	0.57504995184480
0.66642425168314	0.49980912249304	0.37918325617717	0.56634575407519	0.42495004815520

Columns 36 through 40

0.40243603774914	0.29156780227853	0.22559974425105	0.18811982446663	0.16738252919335
0.59756396225086	0.70843219772147	0.77440025574895	0.81188017553337	0.83261747080665

Columns 41 through 45

0.15607679696115 0.34608742211544 0.25755818811001 0.20611815751162 0.17729156119713
0.84392320303885 0.65391257788456 0.74244181188999 0.79388184248838 0.82270843880287

Columns 46 through 50

0.16146433463862 0.35122941115435 0.51396198230507 0.62992433142680 0.43996370882442
0.83853566536138 0.64877058884565 0.48603801769493 0.37007566857320 0.56003629117558

Columns 51 through 55

0.31479212786679 0.23910214112447 0.19568814324264 0.38316120637540 0.27981962271384
0.68520787213321 0.76089785887553 0.80431185675736 0.61683879362460 0.72018037728616

Columns 56 through 60

0.21883130224182 0.1843459801077 0.16531604728747 0.15495663413151 0.14935867950418
0.78116869775818 0.81565440198923 0.83468395271253 0.84504336586849 0.85064132049582

Columns 61 through 65

0.33963038975670 0.50494259194870 0.35613796076880 0.51774617922745 0.36445964204260
0.66036961024330 0.49505740805130 0.64386203923120 0.48225382077255 0.63554035795740

Columns 66 through 70

0.26853542264432 0.44717458257519 0.31930866128682 0.24174706107252 0.19717680140609
0.73146457735568 0.55282541742481 0.68069133871318 0.75825293892748 0.80282319859391

Columns 71 through 75

0.17235757326362 0.15877836559142 0.15142045391762 0.34161739356776 0.25490343592007
0.82764242673638 0.84122163440858 0.84857954608238 0.65838260643224 0.74509656407993

Columns 76 through 80

0.43558408567198 0.31205752562953 0.23750376628465 0.19478949660408 0.17104399909746
0.56441591432802 0.68794247437047 0.76249623371535 0.80521050339592 0.82895600090254

Columns 81 through 85

0.36029392630249 0.52093514441191 0.63449034433520 0.70478337316691 0.49309035978040
0.63970607369751 0.47906485558809 0.36550965566480 0.29521662683309 0.50690964021960

Columns 86 through 90

0.61607330698114 0.69387867258886 0.48520790875271 0.61076914434047 0.42673132012232
0.38392669301886 0.30612132741114 0.51479209124729 0.38923085565953 0.57326867987768

Columns 91 through 95

0.57011299785161 0.39911626009125 0.55007467404878 0.65324605878697 0.71570304334075
0.42988700214839 0.60088373990875 0.44992532595122 0.34675394121303 0.28429695665925

Columns 96 through 100

0.50103389650267 0.35360916541151 0.51579903995398 0.63113019319904 0.70280743332362
0.49896610349733 0.64639083458849 0.48420096004602 0.36886980680096 0.29719256667638

Column 101

0.49165834658560
0.50834165341440

BACKWARD =

Columns 1 through 5

1.00000000000000 0.82870449605679 0.77615133028273 0.82119410978302 0.98195719107088
1.86321296709204 1.45678801051524 1.23252497474508 1.10018468586452 1.00651593239544

Columns 6 through 10

1.32312857590261 0.89815395436986 1.11440578344636 1.56549148241625 1.11423714346307
0.91382819210860 1.06694978345774 0.95371959655207 0.83718781704897 0.92118491554081

Columns 11 through 15

1.52852751467302 1.13854709591087 0.86649307498572 0.95623360086141 1.21813637093047
0.77869410413767 0.87747865083245 1.20021438710924 1.03160417875605 0.90563936669899

Columns 16 through 20

0.85725638732013 1.00083337764454 0.66868529063847 0.60939526869797 0.57866201223388
1.12908542040408 0.99957672874426 1.33497257603503 1.21471700028698 1.15012282787367

Columns 21 through 25

0.56411904424043 0.55991455140596 0.56435672589945 0.57923301412425 0.61052935110414
1.11519859418115 1.09589145279506 1.08443257777670 1.07621113608299 1.06791897617526

Columns 26 through 30

0.67083878671911 0.78442868910949 0.99700547137998 1.39410598163464 2.13551590138071
1.05621949085254 1.03640503006831 1.00050263222130 0.93406585627305 0.81036180904441

Columns 31 through 35

1.50869825243758 1.15134057178196 0.897526793339519 1.01121181952097 0.71361662276789
 0.74537331167680 0.84854383404521 1.16777397582070 0.99141504442326 1.38753907194836

Columns 36 through 40

0.65067374538095 0.63316884828940 0.65405829433022 0.72174759869600 0.86347364814347
 1.23525761704421 1.15097584928461 1.10078039069008 1.06447354482617 1.02744612847620

Columns 41 through 45

1.13626430542648 0.71490915631056 0.78449047612431 0.95401394518987 1.29169257844077
 0.97479901458507 1.15088615588399 1.07476174105628 1.01193951087102 0.93714099651557

Columns 46 through 50

1.93340182592320 1.36552645769919 1.02722937042800 0.76144405930518 0.70976708536812
 0.82026870054691 0.80211241892538 0.97120624170828 1.40605801518764 1.22800656235388

Columns 51 through 55

0.72796860237705 0.82337255408078 1.03436767342679 0.64729770569343 0.63679921518911
 1.12497425378040 1.05550285101740 0.99163838485787 1.21908777135072 1.14111840558351

Columns 56 through 60

0.66617234591348 0.74723628618182 0.91260962148306 1.22883543197414 1.82315696044788
 1.09351621548295 1.05712698646093 1.01730838588717 0.95803816734033 0.85546712383369

Columns 61 through 65

1.26456167226610 0.92239028765730 1.11742070227362 0.78901886759272 0.84250342093738
 0.86393501082930 1.07915940388606 0.93505135120297 1.22650867756311 1.09031864949156

Columns 66 through 70

1.02052504301084 0.66236824290340 0.63105641107529 0.63726485621151 0.68349173837878
 0.99246484208486 1.27310672643654 1.17306946154569 1.11564762951347 1.07773577887931

Columns 71 through 75

0.78833787444916 0.99399135250811 1.38320321491746 0.91186836275887 1.16229688537340
 1.04407890307849 1.00113411637214 0.93162125458553 1.04572918529596 0.94447695008245

Columns 76 through 80

0.78274015329572 0.87293746314007 1.09966234993784 1.55414458956718 2.41925917845039
 1.16766878693098 1.05763682623174 0.96895698310314 0.86594617781012 0.70715482441332

Columns 81 through 85

1.77375195328388 1.43288892108756 1.23841224548454 1.10002918423058 1.44018686139072
 0.56420903177998 0.52927656873466 0.58613879172062 0.76119601855277 0.57181343452115

Columns 86 through 90

1.21992990184995 1.05931925139744 1.34558249829446 1.10664395120795 1.46484173525698
 0.64708642456886 0.86554232020776 0.67427751096035 0.83265757613692 0.65397982778950

Columns 91 through 95

1.18646897510991 1.63860486766608 1.33955531532536 1.15476311276180 0.99843188021377
 0.75270620913184 0.57582878428754 0.58486271248625 0.70844341926270 1.00394766133452

Columns 96 through 100

1.19573262578887 1.66698653607003 1.32179164312530 1.10432156767840 0.91360806041472
 0.80345622781925 0.63512392233936 0.65720881558461 0.82150750766320 1.20430153418304

Column 101

1.000000000000000
 1.000000000000000

SCALE =

Columns 1 through 5

1.000000000000000 0.275000000000000 0.29886363636364 0.31791825095057 0.33107901031544

Columns 6 through 10

0.33928447453533 0.15592175939491 0.32779472271793 0.33729848529632 0.15706061281256

Columns 11 through 15

0.32677743950743 0.16332475736061 0.17856439615458 0.31000482442795 0.32580974607864

Columns 16 through 20

0.16392123061033 0.32094824168628 0.16697222161182 0.31850981581748 0.33146239380702

Columns 21 through 25

0.33951373903609 0.34420885311942 0.34684540350933 0.34829466720327 0.34908195579922

Columns 26 through 30

0.34950689720627	0.34973546509507	0.34985817768106	0.34992399295916	0.34995927309803
Columns 31 through 35				
0.15002182052229	0.16668787797773	0.18126670178186	0.30817853491550	0.17544474651842
Columns 36 through 40				
0.31218312921358	0.32728684669695	0.33698781730063	0.34276002237803	0.34603951535917
Columns 41 through 45				
0.34785402869558	0.15115671973410	0.33221735056490	0.33996365854037	0.34446466121773
Columns 46 through 50				
0.34698698839525	0.15162812928088	0.16823257347601	0.18247167345169	0.30738162100016
Columns 51 through 55				
0.32400317547786	0.33495568881166	0.34157856265161	0.15462271253373	0.32897339444215
Columns 56 through 60				
0.33801578301254	0.34335226105384	0.34636976017406	0.34803484586235	0.34894129451349
Columns 61 through 65				
0.15056888445662	0.16721765910371	0.31831752320449	0.16866207156727	0.31719720931760
Columns 66 through 70				
0.33060978132127	0.16099684948138	0.32337222402467	0.33456049213740	0.34134713215615
Columns 71 through 75				
0.34524702987697	0.34741871233943	0.34860689301075	0.15074928971779	0.33260847806282
Columns 76 through 80				
0.15980405064301	0.32438639250370	0.33519496650742	0.34171842045009	0.34545591904714
Columns 81 through 85				
0.15246634992103	0.16902571855147	0.18308182513604	0.19301790512933	0.30083145484790
Columns 86 through 90				

0.18064540648078 0.19140641436085 0.30178561614847 0.17995569201586 0.30905769987021

Columns 91 through 95

0.17483899051070 0.312615111268798 0.17242267275798 0.18563153397927 0.19465903014386

Columns 96 through 100

0.29987598370768 0.31865953405602 0.16844080197351 0.18263241599597 0.19272389190492

Column 101

0.30100434958418

7. (Local) maximum likelihood estimation

[estS, estT] = hmmtrain(DNAseq,S,T,'Symbols',{'A','C','G','T'})

estS =

0.82882751711803 0.17117248288197
0.29891709552263 0.70108290447737

estT =

0.23147178529797 0.29174573977644 0.19563648364480 0.28114599128079
0.06060765205038 0.34248443779953 0.59690264136317 0.00000526878691