Μαρκοβιανές Αλυσίδες

Στοχαστική Ανέλιξη είναι μια συλλογή τ.μ. $\left\{\theta_t^*,\ t \in T\right\}$. Ο χώρος Τ (συνήθως είναι χρόνος) μπορεί να είναι είτε διακριτός είτε συνεχής και καλείται **παραμετρικός χώρος**. Το σύνολο των δυνατών τιμών των τ.μ. συμβολίζεται συνήθως με S και καλείται **χώρος καταστάσεων** (state space). Ο χώρος καταστάσεων μπορεί να είναι επίσης διακριτός είτε συνεχής. **Μαρκοβιανή Ανέλιξη** είναι μια στοχαστική ανέλιξη στην οποία από όλο το "παρελθόν" της μόνο η πιο πρόσφατη κατάσταση καθορίζει το "μέλλον", δηλαδή:

$$P(\theta_{t+1}^* \in A \mid \theta_0^*, \ldots, \theta_t^*) = P(\theta_{t+1}^* \in A \mid \theta_t^*)$$

Έστω ο S διακριτός (ότι θα πούμε γενικεύεται και σε περίπτωση που είναι συνεχής).

Παράδειγμα

- Έστω ένα αντικείμενο που κινείται στο χώρο {...,-2,-1,0,1,2,...}, ξεκινώντας από το 0. Αν το χρόνο t είναι στη θέση i τότε με πιθανότητα p_2 μένει στην ίδια θέση, με πιθανότητα p_3 πηγαίνει στη θέση (i+1) και με πιθανότητα p_1 πηγαίνει στη θέση (i-1), όπου $p_1+p_2+p_3=1$. Οι παραπάνω διαδικασία είναι προφανώς μια Μαρκοβιανή Αλυσίδα (και καλείται **τυχαίος περίπατος**), και οι πιθανότητες p_1 ,p_2 , p_3 καλούνται πιθανότητες μετάβασης.

Ιδιότητες

- **Μη Υποβιβάσιμη** (Irreducible). Μια Μαρκοβιανή αλυσίδα ονομάζεται μη υποβιβάσιμη αν ανεξαρτήτως της κατάστασης που ξεκινάει θα επισκεφτεί κάθε άλλη κατάσταση σε πεπερασμένο αριθμό επαναλήψεων με θετική πιθανότητα.
- **Απεριοδική** (Aperiodic). Μια Μαρκοβιανή αλυσίδα ονομάζεται απεριοδική όταν δεν υπάρχει ακέραιος d>1:
$$p_{ii}^{n} = P(\theta_n^* = i \mid \theta_0^* = i) = 0 \quad \text{όταν } n \neq kd \quad \text{με } k=1,2,\ldots$$
- **Γνήσια Επαναληπτική** (Positive Recurrent). Μια Μαρκοβιανή αλυσίδα ονομάζεται γνήσια επαναληπτική όταν για κάθε κατάσταση i από την οποία ξεκινά η αλυσίδα, επανέρχεται σε αυτή με πιθανότητα 1 σε πεπερασμένο αναμενόμενο χρόνο.

Ιδιότητες

- Στο προηγούμενο παράδειγμα η αλυσίδα είναι μη υποβιβάσιμη και απεριοδική, αλλά μπορεί να μην είναι γνήσια επαναληπτική (γιατί ο χώρος καταστάσεων δεν είναι φραγμένος, οπότε μπορεί να περιμένεις για πάντα μέχρι να επιστρέψεις από εκεί που άρχισες). Αν S = {-k...,-2,-1,0,1,2,...k}, κρατώντας τις ίδιες πιθανότητες μετάβασης, και απλά απορρίπτοντας κάθε κίνηση εκτός του S, τότε η αλυσίδα είναι και γνήσια επαναληπτική.

Παράδειγμα

```
rw.sim <- function( k, p, theta.start, n.sim, seed ) {
  set.seed( seed )
  theta <- rep( 0, n.sim + 1 )
  theta[ 1 ] <- theta.start
  for ( i in 1:n.sim ) {
    theta[ i + 1 ] <- move( k, p, theta[ i ] )
  }
  return( table( theta ) )
}
```

Παράδειγμα

```
move <- function( k, p, theta ) {

  repeat {

    increment <- sample( x = c( -1, 0, 1 ), size = 1, prob = p )

    theta.next <- theta + increment

    if ( abs( theta.next ) <= k ) {

      return( theta.next )

      break

    }

  }

}
```

Παράδειγμα

```
> p <- c( 1, 1, 1 ) / 3

> k <- 5

> theta.start <- 0

> seed <- c( 6425451, 9626954 )

> rw.sim( k, p, theta.start, 10, seed )

theta
0 1 2
5 5 1

> rw.sim( k, p, theta.start, 100, seed )

-2 -1  0  1  2  3  4  5
 7  9 16 17 23 14  8  7
```

Παράδειγμα

```
> rw.sim( k, p, theta.start, 1000, seed )

 -5  -4  -3  -2  -1   0   1   2   3   4   5
 65 115 123 157 148 123 106  82  46  21  15

> rw.sim( k, p, theta.start, 10000, seed )

  -5   -4   -3   -2   -1    0    1    2    3    4    5
 581  877  941  976  959 1034 1009  982 1002  959  681

> rw.sim( k, p, theta.start, 100000, seed )

  -5   -4   -3   -2   -1    0    1    2    3    4    5
6515 9879 9876 9631 9376 9712 9965 9749 9672 9352 6274

> rw.sim( k, p, theta.start, 1000000, seed )

   -5    -4    -3    -2    -1     0     1     2     3     4     5
65273 98535 97715 96708 95777 96607 96719 96361 96836 95703 63767
```

Παράδειγμα

- Παρατηρούμε ότι η κατανομή των επισκέψεων της αλυσίδας συγκλίνει σε κάτι που μοιάζει με ομοιόμορφη με εξαίρεση τις καταστάσεις -5,+5. Αν συμβολίσουμε με q_1 την οριακή πιθανότητα της επίσκεψης των καταστάσεων {-4,..,+4} και με q_2 την οριακή πιθανότητα της επίσκεψης των καταστάσεων -5,+5, τότε λόγω συμμετρίας μπορούμε να υποθέσουμε ότι $9\,q_1 + 2\,q_2 = 1$ και $q_1 = 3/2\,q_2$, από όπου καταλήγουμε στο ότι $(q_1,\ q_2) = (3/31, 2/31) \approx (0.096774, 0.0645516)$. Οι αντίστοιχες εμπειρικές πιθανότητες είναι

$$\left[\frac{98535+\ldots+95703}{(9)(1000001)}, \frac{65273+63767}{(2)(1000001)}\right] \doteq (0.096773, 0.064520).$$

Παράδειγμα

- Μπορείτε εύκολα να διαπιστώσετε ότι ξεκινώντας από άλλη κατάσταση η οριακή κατανομή δεν αλλάζει. Αντίθετα αν αλλάξετε τις πιθανότητες μετάβασης το αποτέλεσμα είναι διαφορετικό.

```
> p <- c( 0.2, 0.3, 0.5 )

> rw.sim( k, p, 0, 10, seed )

0 1 2 3
1 3 4 3

> rw.sim( k, p, 0, 100, seed )

 0  1  2  3  4  5
 1  3  6 13 30 48

> rw.sim( k, p, 0, 1000, seed )

  0   1   2   3   4   5
  1  18  71 157 336 418

> rw.sim( k, p, 0, 10000, seed )

  -5   -4   -3   -2   -1    0    1    2    3    4    5
   5   16   19   30   28   74  215  583 1344 3470 4217

> rw.sim( k, p, 0, 100000, seed )

   -5    -4    -3    -2    -1     0     1     2     3     4     5
    5    22    53   132   302   834  2204  5502 13489 34460 42998

> rw.sim( k, p, 0, 1000000, seed )
-5     -4     -3     -2     -1      0      1      2      3      4      5
61    198    511   1380   3398   8591  22117  54872 137209 343228 428436
```

Στάσιμη Κατανομή

- Μια γνήσια επαναληπτική και απεριοδική αλυσίδα καλείται **εργοδική** (**ergodic**). Κάθε εργοδική αλυσίδα έχει μια μοναδική **στάσιμη κατανομή** π (**stationary, equilibrium**):

$$\pi(j) = \sum_{i} \pi(i)\mathbf{P}_{ij}(t), \text{ για κάθε κατάσταση } j \text{ και χρόνο } t \geq 0$$

όπου $\mathbf{P}_{ij}(t) = P(\theta^*_t = j \mid \theta^*_{t-1} = i)$ είναι ο **πίνακας μετάβασης** της αλυσίδας.

Η στάσιμη κατανομή χαρακτηρίζει πλήρως την συμπεριφορά όπου η αλυσίδα θα αποκτήσει τελικά, αφού έχει τρέξει για αρκετό διάστημα, ανεξαρτήτως της αρχικής της κατάστασης.

Metropolis et al (1953)

- Έστω παράμετρος **θ** $\in \Theta$ και **y** τα δεδομένα.
- **ΙΔΕΑ**: Προσομοίωσε τιμές από την $p(\boldsymbol{\theta} \mid \mathbf{y})$ δημιουργώντας μια Μαρκοβιανή Αλυσίδα με τις ακόλουθες ιδιότητες:
  - Χώρο καταστάσεων Θ.
  - Στάσιμη κατανομή $p(\boldsymbol{\theta} \mid \mathbf{y})$.
  - Το $p(\boldsymbol{\theta} \mid \mathbf{y})$ να εμφανίζεται στους υπολογισμούς μόνο υπό μορφή πηλίκου $p(\boldsymbol{\theta} \mid \mathbf{y}) / p(\boldsymbol{\theta}' \mid \mathbf{y})$.

Εργοδικό Θεώρημα

- Αν η Μαρκοβιανή Αλυσίδα είναι εργοδική με στάσιμη κατανομή π και f είναι μια οποιαδήποτε πραγματική συνάρτηση με

$E_{\pi}\left[|f(\theta)|\right] < \infty$, τότε με πιθανότητα 1 καθώς $m \to \infty$

$$\frac{1}{m}\sum_{t=1}^{m} f(\theta_t^*) \to E_{\pi}\left[f(\theta)\right] = \sum_i f(i)\pi(i).$$

Με άλλα λόγια αν η $p(\boldsymbol{\theta} \mid \mathbf{y})$ είναι η στάσιμη κατανομή της αλυσίδας, τότε μπορείς να μάθεις πληροφορίες όπως ο εκ των υστέρων μέσος ή τυπική απόκλιση απλά περιμένοντας να επιτευχθεί η στασιμότητα (να έχει συγκλίνει η αλυσίδα). Εν συνεχεία απλά καταγράφεις τις τιμές μετά από αυτό το χρονικό διάστημα.

Ερωτήματα

- Πόσο πρέπει να περιμένεις για να επιτευχθεί η στασιμότητα;
- Πόσο μεγάλο πρέπει να είναι το m (μετά την στασιμότητα για πόσο πρέπει να τρέξεις την αλυσίδα σου);
- Από που να ξεκινήσεις;
- **Απαντήσεις:** Συνήθως ξεκινάμε από μια λογική τιμή και τρέχουμε την αλυσίδα για b επαναλήψεις (**burn in period**) μέχρις ότου επιτευχθεί η στασιμότητα. Εν συνεχεία πετάμε τις b αυτές τιμές και τρέχουμε την αλυσίδα για m ακόμα επαναλήψεις καταγράφοντας τις τιμές. Για να βεβαιωθούμε για όλα τα παραπάνω χρησιμοποιούμε **διαγνωστικούς ελέγχους** (**MCMC diagnostics**).