g-adic representation

Natural number x and base g $\quad g \in \mathbb{N}\setminus\{0,1\}$

Division with smallest remainder: Take n so large that $g^{n+1} > x$ $\quad (\to\infty)$

$$x : g^n = a_n + R \qquad \underbrace{a_n \in \{0, 1, \dots, g-1\}}_{\text{because } g^{n+1} > x}$$

$$\Leftrightarrow \underbrace{(x - R)}_{g^{n+1} > x \geq} = a_n g^n \qquad \text{with } R < g^n$$

$$\Rightarrow \quad x = a_n g^n + R$$

$$R : g^{n-1} = a_{n-1} + R_2 \quad \Leftrightarrow \quad R = a_{n-1} g^{n-1} + R_2$$

$$R_2 < g^{n-1}$$

$$a_{n-1} \in \{0, \dots, g-1\}$$

$$x = a_n g^n + a_{n-1} g^{n-1} + \underbrace{R_2}_{a_{n-2} g^{n-2} + R_3}$$

$$\Rightarrow \quad x = a_n g^n + a_{n-1} g^{n-1} + \dots + a_0 g^0$$

$g = 10 \qquad g^4 = 10\,000 > 1015 > 1000 = g^3$

$1015 : g^3 \qquad 1015 : 1000 = \underline{1} + 15R$

$15 : 100 = 0 + 15R$

$15 : 10 = 1 + 5R$

$5 : 1 = 5$

$1 \cdot 1000 + 0 \cdot 100 + 1 \cdot 10 + 5 \cdot 1$

$1 \cdot \underline{10^3} + 0 \cdot \underline{10^2} + 1 \cdot \underline{10^1} + 5 \cdot \underline{10^0} = (1, 0, 1, 5)_{10}$

$g = 2$

n	2^n
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

$x = 1015$

$2^{10} > 1015 > 2^9$

$1015 : 512 = 1 + 503\,R$

$503 : 256 = 1 + 247\,R$

$247 : 128 = 1 + 119\,R$

$119 : 64 = 1 + 55\,R$

$55 : 32 = 1 + 23\,R$

$23 : 16 = 1 + 7\,R$

$7 : 8 = 0 + 7\,R$

$7 : 4 = 1 + 3\,R$

$3 : 2 = 1 + 1\,R$

$1 : 1 = 1$

$(1,1,1,1,1,1,0,\underset{a_3}{1},\underset{a_2}{1},\underset{a_1}{1},\underset{a_0}{1})$ binary representation of 1015

$20^2 = 400$

$20 \cdot 18 = \underline{\underline{360}}$

Set of alleles per marker $\mathcal{A} = \{A_1, \ldots, A_n\}$

obs. at this marker corresponds to subset θ of $\mathcal{A}$ $\quad \theta \subseteq \mathcal{A}$

$\theta \in \mathcal{P}(\mathcal{A})$

$\sigma \subseteq \mathcal{A}$

$$\{A_1, A_2, \ldots, A_n\}$$
$$\begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 1 & 0 & 1 \end{array}$$

$1 \hat{=} \in$

$0 \hat{=} \notin$

$\sigma \hat{=}$ 0-1 vector of length n $\quad (0, 1, 0, \ldots 1) = \underline{x} \hat{=}$ binary number

$\longleftrightarrow$ binary number

$\langle (2^0, 2^1, 2^2, \ldots 2^n) \mid \underset{\uparrow}{\underline{x}} \rangle$ ... decimal number

binary representation

Eg: $\quad A_1, A_2, A_3$

subset	A_1 ($2^0=1$)	A_2 ($2^1=2$)	A_3 ($2^2=4$)	decimal nrs
$\emptyset$	0	0	0	0
$\{A_1\}$	1	0	0	1
$\{A_2\}$	0	1	0	2
$\{A_1, A_2\}$	1	1	0	3
	0	0	1	4
	1	0	1	5
	0	1	1	6
	1	1	1	$7 = 2^3 - 1$

2^n

We mix bases

3 digits bases $g_1, g_2, g_3 \dots g_n$

For each digit mod $\{0, \dots g_i - 1\}$

$g_1 = 10 \quad g_2 = 5 \quad g_3 = 8$

$$x : (g_1 \cdot g_2 \cdots g_n) = \underbrace{a_n}_{\in\{0, \dots g_n - 1\}} + R$$

$$R : (g_1 \cdot g_2 \cdots g_{n-1}) = a_{n-1} + R_2$$

$$\vdots$$

	Marker 1	Marker 2	Marker 3
# alleles	n_1	n_2	n_3
	$\downarrow$	$\downarrow$	$\downarrow$
# possible observations	$\underbrace{2^{n_1}}_{g_1}$	$\underbrace{2^{n_2}}_{g_2}$	$\underbrace{2^{n_3}}_{g_3}$
obs in binary representation	x_1	x_2	x_3

$$\Rightarrow \boxed{x_1 + x_2 \cdot g_1 + x_3 \cdot g_1 g_2} = X \; \dots \text{compact representation}$$

Dataset!

	compact representation
Sample 1	X_1
sample 2	X_2
sample 3	X_3
	$\vdots$

\# of alleles $\quad n_1 \quad n_2 \quad n_3, \dots \; n_l$

$$2^{n_1} \quad 2^{n_2} \quad 2^{n_3} \quad \dots \quad 2^{n_l}$$

$$1 \quad g_1 \quad g_2 \quad g_3 \quad \dots \quad g_l$$

$$(1, g_1, g_2, g_3, \dots g_{l-1})$$

$$\xrightarrow{\text{cu prod}} (1, g_1, g_1 \cdot g_2, g_1 \cdot g_2 \cdot g_3, \dots\dots (g_1 \cdot g_2 \cdots g_{l-1}))$$

$$x \to f(x) \to f(f(x)) \to f(f(f(x))) \to \dots$$

$$x \to f(x, v_1) \to f(f(x), v_2) \to f(f(f(x)), v_3) \to \dots$$

$$\underset{\downarrow\; a_1}{x : f_1 = a_1 + r_1} \to \underset{\downarrow\; a_2}{r_1 : f_2 = a_2 + r_2} \to \underset{\downarrow\; a_3}{r_2 : f_3 = a_3 + r_3} \to$$